

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:16:49 ; Search time 7.30435 seconds  
(without alignments)  
530.147 Million cell updates/sec

Title: US-09-856-319B-4\_COPY\_34\_264

Perfect score: 1252

Sequence: 1 IVNGENAVPGSWPQVSLQD.....AMVTRVSKFTWVNOVMAYN 231

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 106657 seqs, 16763532 residues

Total number of hits satisfying chosen parameters: 106657

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB pep:\*

2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB pep:\*

3: /cgn2\_6/ptodata/1/pubpaa/US05\_NEW\_PUB pep:\*

4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB pep:\*

5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB pep:\*

6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB pep:\*

7: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB pep:\*

8: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB pep:\*

9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB pep:\*

10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB pep:\*

11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB pep:\*

12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB pep:\*

13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB pep:\*

14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	756	60.4	192	10	US-09-925-297-529
2	755.5	60.3	263	10	US-09-888-615-96
3	464.5	37.1	251	9	US-09-961-721-4
4	464.5	37.1	791	9	US-09-967-386-1
5	464.5	37.1	810	10	US-09-946-893-2
6	462	36.9	812	9	US-09-333-325-1
7	462	36.9	812	10	US-09-788-142-1
8	462	36.9	812	10	US-09-761-120-1
9	462	36.9	812	10	US-09-873-676-81
10	459.5	36.7	270	10	US-09-923-779-152
11	457.5	36.5	249	9	US-09-961-721-5
12	457.5	36.5	273	10	US-09-925-297-529
13	455.5	36.4	453	9	US-09-978-285A-69
14	455.5	36.4	453	9	US-09-978-697-69
15	455.5	36.4	453	9	US-09-978-192A-69
16	455.5	36.4	453	12	US-10-052-586-64
17	455	36.3	248	10	US-09-925-301-1017
18	455	36.3	327	10	US-09-804-156-16
19	455	36.3	327	10	US-09-946-633-8

20	455	36.3	454	10	US-09-888-615-103	Sequence 103, App
21	448.5	35.8	343	10	US-09-948-094-2	Sequence 2, Appl
22	438.5	35.0	226	9	US-09-961-721-6	Sequence 6, Appl
23	438.5	35.0	226	9	US-10-045-367A-5	Sequence 5, Appl
24	438.5	35.0	226	10	US-09-910-151-6	Sequence 6, Appl
25	430.5	34.4	393	9	US-10-012-896-934	Sequence 934, App
26	430.5	34.4	393	10	US-09-759-143-934	Sequence 934, App
27	430.5	34.4	393	10	US-09-780-669-934	Sequence 934, App
28	430.5	34.4	393	10	US-09-822-827-934	Sequence 934, App
29	430.5	34.4	492	9	US-10-012-896-932	Sequence 932, App
30	430.5	34.4	492	10	US-09-759-143-932	Sequence 932, App
31	430.5	34.4	492	10	US-09-780-669-932	Sequence 932, App
32	430.5	34.4	492	10	US-09-822-827-932	Sequence 932, App
33	428.5	34.2	492	9	US-10-012-896-895	Sequence 895, App
34	428.5	34.2	492	10	US-09-759-143-895	Sequence 895, App
35	428.5	34.2	492	10	US-09-780-669-895	Sequence 895, App
36	428.5	34.2	492	10	US-09-822-827-895	Sequence 895, App
37	428.5	34.2	492	10	US-09-822-827-895	Sequence 14, Appl
38	427.5	34.1	384	9	US-09-981-353-23	Sequence 23, Appl
39	424	33.9	320	10	US-09-888-615-90	Sequence 90, Appl
40	423.5	33.8	457	10	US-09-888-615-110	Sequence 110, App
41	423	33.8	223	10	US-09-910-071-14	Sequence 14, Appl
42	422.5	33.7	414	10	US-09-820-893-69	Sequence 69, Appl
43	422.5	33.7	480	10	US-09-820-893-108	Sequence 108, App
44	422	33.7	1169	9	US-09-870-759-126	Sequence 126, App
45	421.5	33.7	283	10	US-09-988-975A-1	Sequence 1, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-925-297-529

; Sequence 529, Application US/09925297

; Patent No. US20020081659A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

; FILE REFERENCE: PA105

; CURRENT APPLICATION NUMBER: US/09/925,297

; CURRENT FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/05989

; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270

; PRIOR FILING DATE: 1999-03-12

; NUMBER OF SEQ ID NOS: 928

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 529

; LENGTH: 192

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-925-297-529

Query Match

Best Local Similarity 60.4%; Score 756; DB 10; Length 192;

Matches 135; Conservative 16; Mismatches 9; Indels 0; Gaps 0;

QY 1 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVTAACQVTPGRHFVVLGEYDRS 60

||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 30 IVNGENAVLGSWPQVSLQDSSGFFHFCGSLISQSWVTAACQVTPGRHFVVLGEYDRS 89

||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 61 SNAEPQVLSIARAIHPNNANTMNDLTLKLASPARVTAQVSPVCLASTNEALPSGL 120

||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 90 SNAEPLOVLSVRAIHPNSWNTMNDVTLKLASPAQYTRISPVCCLASSNEALTEGL 149

||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 121 TCVTGWRISGVNGVTPARLQOVLPVTVNOCROYGA 160

||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 150 TCVTGWRISGVNGVTPARLQOVLPVTVNOCROYWS 189

||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

##### RESULT 2

US-09-888-615-96

; Sequence 96, Application US/09888615

```
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: FLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR FILING DATE: 2001-06-26
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 96
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-96

Query Match      60.3%; Score 755.5; DB 10; Length 263;
Best Local Similarity 58.4%; Pred. No. 7.3e-62;
Matches 135; Conservative 35; Mismatches 60; Indels 1; Gaps 1;

Qy 1 IVNGENAVPGSWPQVSLQDNTGFHFCGGLISPNWVYTAACOVTPCRHFVVLGEYDRS 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 34 IVNGEDAVPGSWPQVSLQDNTGFHFCGGLISPDWVYTAACGVRTS-DVVVAGEFDQG 92
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 61 SNAEPQVLSIARAITHPNMNTMNDLTLKLLASPARYTAQVSPVCLASTNEALPISGL 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 93 SDEENIQVLKAKVFNKPSILTVNNDITLLKLTAPARETSQTVSAVCLPSADDDFPAGT 152
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 121 TCVTTGGRISGVGNVTPARLQOVVPLVTVNQCROYNGARITDAMICAGSGAGSCQGD 180
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 153 LCATTGKTKYNKATPDKLQQAALPLLSNAECKSKMGRRTDVMICAGASGVSSCMGD 212
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 181 SGGPLVCQKGTWLVIGSVNGTKNCNIQAPAMYTRVSKFSTWVNOVMAYN 231
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 213 SGGPLVCQKDGAWTLVGIVSGSRTCTTTPAVYARVTKLIPWQKILAN 263
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 3
US-09-961-721-4
; Sequence 4, Application US/09961721
; Patent No. US20020156005A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: m32404, A NOVEL HUMAN TRYPSIN AND USES
; FILE REFERENCE: THEREOF
; FILE REFERENCE: 10448-096001
; CURRENT APPLICATION NUMBER: US/09/961,721
; CURRENT FILING DATE: 2001-09-24
; PRIOR FILING DATE: 2000-09-24
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus sequence
US-09-961-721-4

Query Match      37.1%; Score 464.5; DB 9; Length 251;
Best Local Similarity 44.2%; Pred. No. 2.5e-35;
Matches 111; Conservative 26; Mismatches 79; Indels 35; Gaps 12;

Qy 9 PGSW--PQVSLQDNTG----FHCGLISPNWVYTAAC--QVTPGRHF-----VV 53
    ||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 PGFSGSPWQVSLQVRSRGGSRKHFCCGGLISENVLTAAHCVCASAPASSVRVLSVR 60
    ||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: FLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR FILING DATE: 2001-06-26
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 96
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-96

Query Match      60.3%; Score 755.5; DB 10; Length 263;
Best Local Similarity 58.4%; Pred. No. 7.3e-62;
Matches 135; Conservative 35; Mismatches 60; Indels 1; Gaps 1;

Qy 1 IVNGENAVPGSWPQVSLQDNTGFHFCGGLISPNWVYTAACOVTPCRHFVVLGEYDRS 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 34 IVNGEDAVPGSWPQVSLQDNTGFHFCGGLISPDWVYTAACGVRTS-DVVVAGEFDQG 92
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 61 SNAEPQVLSIARAITHPNMNTMNDLTLKLLASPARYTAQVSPVCLASTNEALPISGL 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 93 SDEENIQVLKAKVFNKPSILTVNNDITLLKLTAPARETSQTVSAVCLPSADDDFPAGT 152
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 121 TCVTTGGRISGVGNVTPARLQOVVPLVTVNQCROYNGARITDAMICAGSGAGSCQGD 180
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 153 LCATTGKTKYNKATPDKLQQAALPLLSNAECKSKMGRRTDVMICAGASGVSSCMGD 212
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 181 SGGPLVCQKGTWLVIGSVNGTKNCNIQAPAMYTRVSKFSTWVNOVMAYN 231
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 213 SGGPLVCQKDGAWTLVGIVSGSRTCTTTPAVYARVTKLIPWQKILAN 263
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 3
US-09-961-721-4
; Sequence 4, Application US/09961721
; Patent No. US20020156005A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: m32404, A NOVEL HUMAN TRYPSIN AND USES
; FILE REFERENCE: THEREOF
; FILE REFERENCE: 10448-096001
; CURRENT APPLICATION NUMBER: US/09/961,721
; CURRENT FILING DATE: 2001-09-24
; PRIOR FILING DATE: 2000-09-24
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus sequence
US-09-961-721-4

Query Match      37.1%; Score 464.5; DB 9; Length 251;
Best Local Similarity 44.2%; Pred. No. 2.5e-35;
Matches 111; Conservative 26; Mismatches 79; Indels 35; Gaps 12;

Qy 9 PGSW--PQVSLQDNTG----FHCGLISPNWVYTAAC--QVTPGRHF-----VV 53
    ||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 PGFSGSPWQVSLQVRSRGGSRKHFCCGGLISENVLTAAHCVCASAPASSVRVLSVR 60
    ||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

; Patent No. US20020159992A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Henkin, Jack
; APPLICANT: Davidson, Donald J.
; TITLE OF INVENTION: ANTANGIOGENIC POLYPEPTIDES AND METHODS
; FILE REFERENCE: 6738-US-02
; CURRENT APPLICATION NUMBER: US/09/967,386
; CURRENT FILING DATE: 2001-09-26
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 791
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-967-386-1

Query Match      37.1%; Score 464.5; DB 9; Length 791;
Best Local Similarity 41.8%; Pred. No. 1e-34;
Matches 100; Conservative 29; Mismatches 93; Indels 17; Gaps 7;

Qy 1 IVNGENAVPGSWPQVSLQDNTGFHFCGGLISPNWVYTAAC--QVTPGRHFVWLGEY 57
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 562 VVGCVAHPHSWPQVSLRTRFGMHFCGGLISPEWVLTAAHLEKSPRSPSYKVILGAH 621
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 58 DRSSNAEP-VQVLSIARAITHPNMNTMNDLTLKLLASPARYTAQVSPVCLASTNEAL 116
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 622 -QEVNLEPHVQIEVSRLEPT-----RKDIALLLKSSPAVITDKVIPACLPSPV 674
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 117 PSGLTCTVTGGRISGVGNVTPARLQOVVPLVTVNQCROY--NGARITDAMICAG--GS 172
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 675 ADRTCEFTIGWETQGTFGA--GLLKEAQLPVIENKVCNRYEFLNGRVQSTELCAGHLAG 732
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 173 GASSCOGDSGGPLVCOKGNTWLVIGIVSWGTKNCNIQAPAMYTRVSKFSTWVNOVMAYN 231
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 733 GTDSCQGDGGPLVCFFERDKYIILOQGVTSWGLGCRPNKPGVYVRSRFTVWIEGVMRNN 791
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 5
US-09-946-893-2
; Sequence 2, Application US/09946893
; Patent No. US20020072494A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yihai
; TITLE OF INVENTION: Materials and methods relating to endothelial cell growth
; FILE REFERENCE: Mawburn
; CURRENT APPLICATION NUMBER: US/09/946,893
; CURRENT FILING DATE: 2001-09-05
```







```
Best Local Similarity 42.3%; Pred. No. 7.8e-35;
Matches 102; Conservative 38; Mismatches 88; Indels 13; Gaps 9;

QY 1 IVNGENAVPGSWPQVLSQ-DNTG--FHFCGGLSPNNVVTAAHCQVTPGRHFVVLGEY 57
Db 29 VVHGEDAVPYSWPQVLSQYKESGFYHTCGGSLIAPDWMVVTAGHCISRDLTYQVVLGEY 88
QY 58 DRSSNAEPVQVLSI--ARAITHPNNTM--NNDLTLLKLASPARYTAQVSPVCLASTN 113
Db 89 NLAVKEGPEQVPIINSEELFVHPLNRSVCAGNDIALIKLSRQALGDVQLASLPAG 148
QY 114 EALPSGLTCVTGWRISGVGNVTPARLQVVLPLVTNQCQ--YWGARITDAMICAGG 171
Db 149 DILPNKTPCYITGWRGLYNGPL-PDKLQARLPVVDYKHCRRWNWGSTVKKTWVCAGG 207
QY 172 SGASSCGDSGGPLVC-QKGNVTWVLIGVSWGK-NCN-IOAPAMYTRVSKFSTWINQVM 228
Db 208 YIRSGCNGDSGGPLNCPTEDEGQVHGVTFSVSGFCNFINKPTVTRVSAFIDWIEETI 267
QY 229 A 229
Db 268 A 268

RESULT 11
US-09-961-721-5
; Sequence 5, Application US/09961721
; Patent No. US20020156005A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: m32404, A NOVEL HUMAN TRYPSIN AND USES
; FILE REFERENCE: 10448-096001
; CURRENT APPLICATION NUMBER: US/09/961,721
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 60/235,023
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus sequence
US-09-961-721-5

Query Match 36.5%; Score 457.5; DB 9; Length 249;
Best Local Similarity 44.1%; Pred. No. 1.1e-34;
Matches 108; Conservative 25; Mismatches 79; Indels 33; Gaps 11;

QY 13 PQVSLQDNTG---FHFCGGLSPNNVVTAAHC-----QVTPGRHF-----VVLGEYDR 59
Db 5 PQVSLQVRSRGGGRKHFHFCGGLSPNNVVTAAHC-----NNDLTLLKLASPARYTAQ-VSPVCLAST 64
QY 60 SSNAEPVQVLSIARA-ITHPNNTM-----NNDLTLLKLASPARYTAQ-VSPVCLAST 112
Db 65 SUTGTEQKDFVKKTIIVHPNTNPOTLDNGAYDNDIALLKSPGVTGLDTRVPCLPSPA 124
QY 113 NEALPSGLTCVTGWRISGVGNVTPARLQVVLPLVTNQCQ--QYWGAR-----ITD 164
Db 125 SSDLPVGTCTVSGWRRTKGLSLDYLQEVVVPVSRSETCRSAYEYGGTDKVEFVTD 184
QY 165 AMICAGG-SGASSCGDSGGPLVCQKGN---TWVLIGVSWGKNC-NIOAPAMYTRVSK 219
Db 185 NMICAGLGKDACQDSGGPLVCSDGNDGRWELVGIWGSYGCGARGNKPQVYTRVSS 244
QY 220 FSTWI 224
Db 245 YLDWI 249

RESULT 12
US-09-925-297-695
; Sequence 695, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 695
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (27)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (28)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (34)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-695

Query Match 36.5%; Score 457.5; DB 10; Length 273;
Best Local Similarity 42.3%; Pred. No. 1.2e-34;
Matches 102; Conservative 37; Mismatches 89; Indels 13; Gaps 9;

QY 1 IVNGENAVPGSWPQVLSQ-DNTG--FHFCGGLSPNNVVTAAHCQVTPGRHFVVLGEY 57
Db 32 VVHGEDAVPYSWPQVLSQYKESGFYHTCGGSLIAPDWMVVTAGHCISRDLTYQVVLGEY 91
QY 58 DRSSNAEPVQVLSI--ARAITHPNNTM--NNDLTLLKLASPARYTAQVSPVCLASTN 113
Db 92 NLAVKEGPEQVPIINSEELFVHPLNRSVCAGNDIALIKLSRQALGDVQLASLPAG 151
QY 114 EALPSGLTCVTGWRISGVGNVTPARLQVVLPLVTNQCQ--YWGARITDAMICAGG 171
Db 152 DILPNKTPCYITGWRGLYNGPL-PDKLQARLPVVDYKHCRRWNWGSTVKKTWVCAGG 210
QY 172 SGASSCGDSGGPLVC-QKGNVTWVLIGVSWGK-NCN-IOAPAMYTRVSKFSTWINQVM 228
Db 211 YIRSGCNGDSGGPLNCPTEDEGQVHGVTFSVSGFCNFINKPTVTRVSAFIDWIEETI 270
QY 229 A 229
Db 271 A 271

RESULT 13
US-09-978-295A-69
; Sequence 69, Application US/09978295A
; Patent No. US20020156006A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
```

APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James;  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
TITLE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2630PIC11  
CURRENT APPLICATION NUMBER: US/09/978,295A  
CURRENT FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083742  
PRIOR FILING DATE: 1998-04-30  
PRIOR APPLICATION NUMBER: 60/084366  
PRIOR FILING DATE: 1998-05-05

:	PRIOR APPLICATION NUMBER:	60/084414
:	PRIOR FILING DATE:	1998-05-06
:	PRIOR APPLICATION NUMBER:	60/084441
:	PRIOR FILING DATE:	1998-05-06
:	PRIOR APPLICATION NUMBER:	60/084637
:	PRIOR FILING DATE:	1998-05-07
:	PRIOR APPLICATION NUMBER:	60/084639
:	PRIOR FILING DATE:	1998-05-07
:	PRIOR APPLICATION NUMBER:	60/084640
:	PRIOR FILING DATE:	1998-05-07
:	PRIOR APPLICATION NUMBER:	60/084598
:	PRIOR FILING DATE:	1998-05-07
:	PRIOR APPLICATION NUMBER:	60/084600
:	PRIOR FILING DATE:	1998-05-07
:	PRIOR APPLICATION NUMBER:	60/084627
:	PRIOR FILING DATE:	1998-05-07
:	PRIOR APPLICATION NUMBER:	60/084643
:	PRIOR FILING DATE:	1998-05-07
:	PRIOR APPLICATION NUMBER:	60/085339
:	PRIOR FILING DATE:	1998-05-13
:	PRIOR APPLICATION NUMBER:	60/085338
:	PRIOR FILING DATE:	1998-05-13
:	PRIOR APPLICATION NUMBER:	60/085323
:	PRIOR FILING DATE:	1998-05-13
:	PRIOR APPLICATION NUMBER:	60/085582
:	PRIOR FILING DATE:	1998-05-15
:	PRIOR APPLICATION NUMBER:	60/085700
:	PRIOR FILING DATE:	1998-05-15
:	PRIOR APPLICATION NUMBER:	60/085689
:	PRIOR FILING DATE:	1998-05-15
:	PRIOR APPLICATION NUMBER:	60/085579
:	PRIOR FILING DATE:	1998-05-15
:	PRIOR APPLICATION NUMBER:	60/085580
:	PRIOR FILING DATE:	1998-05-15
:	PRIOR APPLICATION NUMBER:	60/085573
:	PRIOR FILING DATE:	1998-05-15
:	PRIOR APPLICATION NUMBER:	60/085704
:	PRIOR FILING DATE:	1998-05-15
:	PRIOR APPLICATION NUMBER:	60/085697

;  
; PRIOR APPLICATION NUMBER: 60/080105  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080107  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080165  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080194  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080327  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080328  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080333  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080334  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/081070  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081049  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081071  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081195  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081203  
; PRIOR FILING DATE: 1998-04-09  
; PRIOR APPLICATION NUMBER: 60/081229  
; PRIOR FILING DATE: 1998-04-09  
; PRIOR APPLICATION NUMBER: 60/081955  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081817  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081819  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081952  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081838  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/082568  
; PRIOR FILING DATE: 1998-04-21  
; PRIOR APPLICATION NUMBER: 60/082569  
; PRIOR FILING DATE: 1998-04-21  
; PRIOR APPLICATION NUMBER: 60/082704  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/082804  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/082700  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/082797  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/082796  
; PRIOR FILING DATE: 1998-04-23  
; PRIOR APPLICATION NUMBER: 60/083336  
; PRIOR FILING DATE: 1998-04-27  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/083392  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083495  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083496  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083499  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083545  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083554  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083558  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083559  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083500

;  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083742  
; PRIOR FILING DATE: 1998-04-30  
; PRIOR APPLICATION NUMBER: 60/084366  
; PRIOR FILING DATE: 1998-05-05  
; PRIOR APPLICATION NUMBER: 60/084414  
; PRIOR FILING DATE: 1998-05-06  
; PRIOR APPLICATION NUMBER: 60/084441  
; PRIOR FILING DATE: 1998-05-06  
; PRIOR APPLICATION NUMBER: 60/084637  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084639  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084640  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084598  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084627  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084643  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/085339  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: 60/085338  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: 60/085323  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: 60/085582  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085700  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085689  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085579  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085580  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085573  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085704  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 36.4%; Score 455.5; DB 9; Length 453;  
Best Local Similarity 37.9%; Pred. No. 3.4e-34;  
Matches 89; Conservative 39; Mismatches 96; Indels 11; Gaps 6;

QY 1 IVNGENAVPGSNPWQVSLQDNTGFHFCGSLISPNWVYTAHC---QVTPGRHFVVLGEY 57  
Db 217 IVGGNMSLLSQWPQASLQFQ-GYHLCGGSVITPLIITAAHCYDLYLPKSWTIOVGLV 275  
QY 58 DRSSNAEPQVLSIARAITHPNNANTMNDLTLKLASPARTAQVSPVCLASTNEALP 117  
Db 276 SLDDNPAPSHL--VEKIVYHSKYKPKRLGALMKGALPLTNEIMIQPVCLNSEENFP 333  
QY 118 SGLTCVTGTGWRISGVGNVTPARLQOVPLVTVNOC--ROYWGARITDAMICAG--GSG 173  
Db 334 DGKVCWTSWGATEDGDASPV-LNHAAYPLISNKNHEDYVGGIISPSMLCAGYLTGG 392  
QY 174 ASSCQDGGGPLVCQKGNWTWVLIGIVSWGPKNCNIQAPAMYTRVSKFSTWINQVM 228  
Db 393 VDSQCGDSGGPLVCQERRLWKLVGATSFYGICAEVKNKPGVYTRVTSFLDWIHEQM 447

RESULT 15  
US-09-978-192A-69  
; Sequence 69, Application US/09978192A  
; Patent No. US20020177553A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Baker Kevin P.

APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan  
APPLICANT: Ferrara, Napoleon  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2630PIC9  
CURRENT APPLICATION NUMBER: US/09/978,192A  
CURRENT FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29

; PRIOR APPLICATION NUMBER: 60/083558  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083559  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083500  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083742  
 ; PRIOR FILING DATE: 1998-04-30  
 ; PRIOR APPLICATION NUMBER: 60/084366  
 ; PRIOR FILING DATE: 1998-05-05  
 ; PRIOR APPLICATION NUMBER: 60/084414  
 ; PRIOR FILING DATE: 1998-05-06  
 ; PRIOR APPLICATION NUMBER: 60/084441  
 ; PRIOR FILING DATE: 1998-05-06  
 ; PRIOR APPLICATION NUMBER: 60/084637  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084639  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084640  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084627  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084643  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/085339  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085338  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085323  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085582  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085700  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085689  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085579  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085580  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085573  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085704  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085697

Query Match 36.4%; Score 455.5; DB 9; Length 453;  
 Best Local Similarity 37.9%; Pred. No. 3.4e-34;  
 Matches 89; Conservative 39; Mismatches 96; Indels 11; Gaps 6;  
 Qy 1 IVNGENAVPGSWPQVSLQDNTGFCGSLISPNWVYTAHC---QVTPGRHFVVLGEY 57  
 Db 217 IVNGMSLLSQWPQASLQFQ-GYHLCGGSVITPLWITTAACHVVDLYLPKSWTIQVGLV 275  
 Qy 58 DRSSNAEPQVLSLSTARITHPNWNTMNDLTLKLASPARYTAQVSPVCLASTNEALP 117  
 Db 276 SLLDNPAPSHL--VEKIVYHSKYKPKRLGNDIALMKLAGPLTFNEMIQVCLPNSSENF 333  
 Qy 118 SGLTCVTTGGRISGVNVTAPRLQVVLPLVTVNOC--ROYMGARITDAMICAG--GSG 173  
 Db 334 DGKVCWTSGWGEDGDSFV-LNHAAPVLSIKNICNHRDVGGLISPSMLCAGYLTGG 392  
 Qy 174 ASSCQDGGPLVCQKQNTWLVIGIVSWGKNCNIPAMYTRVSKFSTWNOVM 228  
 Db 393 VDSQDGGPLVCQERRMLKLVGATSGIGCAEVNPKPGVTVRTVSFLDWIHEQM 447

Search completed: December 20, 2002, 15:21:08  
 Job time : 8.30435 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:14:53 ; Search time 13.087 Seconds  
(without alignments)  
519.349 Million cell updates/sec

Title: US-09-856-319B-2\_COPY\_1\_231

Perfect score: 1221

Sequence: 1 MLLSLTSLVLLGSSWGCG.....GDSGGLVCKGNTWVLIGI 231

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*

2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*

3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*

4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*

5: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep.\*

6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	621.5	50.9	231	2	US-09-027-337-6
2	621.5	50.9	231	4	US-09-644-600-6
3	616.5	50.5	230	4	US-08-944-483-62
4	609	49.9	229	2	US-08-557-146-13
5	609	49.9	229	2	US-09-154-344-13
6	607.5	49.8	228	1	US-08-278-091-10
7	607.5	49.8	228	1	US-08-483-859-10
8	607.5	49.8	228	1	US-08-472-173-10
9	607.5	49.8	228	2	US-08-487-167-10
10	607.5	49.8	228	2	US-08-482-816-10
11	607.5	49.8	228	2	US-08-296-149-10
12	607.5	49.8	228	2	US-08-801-499-10
13	607.5	49.8	228	2	US-08-615-271-10
14	607.5	49.8	228	3	US-09-074-660-10
15	607.5	49.8	228	3	US-09-074-659-10
16	607.5	49.8	228	3	US-09-106-468-10
17	607.5	49.8	228	4	US-09-106-466A-10
18	607.5	49.8	228	4	US-09-106-467-10
19	424	34.7	319	4	US-09-386-642-12
20	419	34.3	241	4	US-08-944-483-59
21	416.5	34.1	242	4	US-08-944-483-57
22	408.5	33.5	454	4	US-09-518-046-2
23	402	32.9	814	1	US-08-750-711-1
24	399	32.7	241	4	US-08-944-483-60
25	398	32.6	791	1	US-08-643-219-1
26	398	32.6	791	3	US-08-851-350-1
27	397.5	32.6	299	4	US-08-944-483-66

28	397	32.5	546	6	5200340-6
29	397	32.5	790	1	US-08-469-486-54
30	397	32.5	790	2	US-08-469-486-54
31	397	32.5	791	2	US-09-131-995-1
32	397	32.5	791	2	US-08-832-087B-1
33	397	32.5	791	4	US-09-132-154-1
34	397	32.5	810	1	US-07-854-603-2
35	397	32.5	810	1	US-08-147-000B-29
36	397	32.5	810	4	US-09-086-514-1
37	397	32.5	810	6	5200340-8
38	395.5	32.4	798	1	US-08-200-900A-2
39	395.5	32.4	798	5	PTC-US94-00616-2
40	392.5	32.1	242	4	US-08-944-483-58
41	390.5	32.0	328	4	US-09-386-642-11
42	389.5	31.9	812	1	US-08-248-629A-1
43	389.5	31.9	812	1	US-08-451-932-1
44	389.5	31.9	812	1	US-08-452-260-1
45	389.5	31.9	812	1	US-08-326-785-1

## ALIGNMENTS

RESULT 1  
US-09-027-337-6

Sequence 6, Application US/09027337B

Patent No. 5972616

GENERAL INFORMATION:

APPLICANT: O'Brien, Timothy J.

APPLICANT: Tanimoto, Hirotooshi

TITLE OF INVENTION: TAGD-15: An Extracellular Serine Protease Overexpressed in

FILE REFERENCE: D6064

CURRENT APPLICATION NUMBER: US/09/027.337B

CURRENT FILING DATE: 1998-02-20

NUMBER OF SEQ ID NOS: 13

SEQ ID NO 6

LENGTH: 231

TYPE: PRT

ORGANISM: Unknown

FEATURE:

OTHER INFORMATION: Serine protease catalytic domain of chymotrypsin (Chymb)

OTHER INFORMATION: homologous to similar domain in TAGD-15

US-09-027-337-6

Query Match 50.9%; Score 621.5; DB 2; Length 231;

Best Local Similarity 56.8%; Pred. No. 5e-53;

Matches 113; Conservative 31; Mismatches 54; Indels 1; Gaps 1;

Qy 33 RIVGENAVLGSWPQVSLQDSSGFFHFCGSLISQSWVYVTAHCNVSFGRHFVVLGEYDR 92

Db 1 RIVGEDAVPCSWPQVSLQDKTFHFCGSLISDQVWVYVTAHCNVRTS-DVVVAGEFDQ 59

Qy 93 SSNAEPQLQVLSVRAITHPSWNSTMMNDVTLKLKASPAQYTTTRISPVCLASSNEALTEG 152

Db 60 GSDEENIQLVIAKFKKPKFSILTNNITLLKATPAFQSQTVAVCLPSADDDPPAG 119

Qy 153 LTCVTITGRLSGVGNVTPAHLLQVVALPLTVNQCQVWDSITDSMTICAGAGASSCQG 212

Db 120 TLCATTGKTKYNAKTPDKLQAAALPLLSNAECKSGRRITDVMICAGASGVSCMG 179

Qy 213 DSGGPLVCOKGNTWVLIGI 231

Db 180 DSGGPLVCOKGNTWVLIGI 198

RESULT 2

US-09-644-600-6

Sequence 6, Application US/09644600

Patent No. 6451500

GENERAL INFORMATION:

APPLICANT: O'Brien, Timothy J.

APPLICANT: Tanimoto, Hirotooshi

;; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease  
;; FILE REFERENCE: Overexpressed in Carcinomas  
;; CURRENT APPLICATION NUMBER: US/09/644,600

;; CURRENT FILING DATE: 2000-08-23  
;; PRIOR APPLICATION NUMBER: 09/421,213

;; PRIOR FILING DATE: 1999-10-20  
;; PRIOR APPLICATION NUMBER: 09/027,337

;; PRIOR FILING DATE: 1998-02-20  
;; NUMBER OF SEQ ID NOS: 98

;; SEQ ID NO 6  
;; LENGTH: 231

;; TYPE: PRT  
;; ORGANISM: Homo sapiens

;; FEATURE:  
;; OTHER INFORMATION: Chymotrypsin

US-09-644-600-6

Query Match 50.9%; Score 621.5; DB 4; Length 231;

Best Local Similarity 56.8%; Pred. No. 5e-53;

Matches 113; Conservative 31; Mismatches 54; Indels 1; Gaps 1;

QY 33 RIVNGENAVLGSNPMQVSLQDSSGFHFCGSLISQSWVYTAHCNVSPGRHFVVLGEYDR 92

DB 1 RIVNGEDAVPGSNPMQVSLQDKTGFHFCGSLISEDWVYVTAHCGVRTS-DVVVAGEFDQ 59

QY 93 SSNAEPLQVLSVSRATTHPSNMTNNNDVTLKLASPAQYTRISPCVCLASSNEALTEG 152

DB 60 GSDEENIQVLKTAKEFKNFPSILTVNNDITLLKATPARFQTSVAVCLPADDFFAG 119

QY 153 LTCVTTGWRGLSGVGNVTPAHQQVALPLVTNQCQRYWDSITDSMICAGGAGSSCOG 212

DB 120 TLCATTGWGKTKYANKTPDKLQQAALPLLSNAECKKSGRRITDVMICAGASGVSSCMG 179

QY 213 DSGGPLVCQKGNWTWLVGI 231

DB 180 DSGGPLVCQKGNWTWLVGI 198

#### RESULT 3

US-08-944-483-62

;; Sequence 62, Application US/08944483

;; Patent No. 6232456

;; GENERAL INFORMATION:

;; APPLICANT: COHEN, MAURICE

;; APPLICANT: COLPITTS, TRACEY L.

;; APPLICANT: FRIEDMAN, PAULA N.

;; APPLICANT: GRANADOS, EDWARD N.

;; APPLICANT: KLASS, MICHAEL R.

;; APPLICANT: RUSSELL, JOHN C.

;; APPLICANT: STEWART, KENT D.

;; APPLICANT: STROUPE, STEVEN D.

;; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS

;; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES

;; NUMBER OF SEQUENCES: 76

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: Abbott Laboratories

;; STREET: 100 Abbott Park Road

;; CITY: Abbott Park

;; STATE: IL

;; COUNTRY: USA

;; ZIP: 60064-3500

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Diskette

;; COMPUTER: IBM Compatible

;; OPERATING SYSTEM: DOS

;; SOFTWARE: FastSeq for Windows Version 2.0

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/944,483

;; FILING DATE:

;; CLASSIFICATION: 424

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER:

;; FILING DATE:

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Becker, Cheryl L.

;; REGISTRATION NUMBER: 35,441

;; REFERENCE/DOCKET NUMBER: 6183.US.01

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: 847/935-1729

;; TELEFAX: 847/938-2623

;; TELEX:

;; INFORMATION FOR SEQ ID NO: 62:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 230 amino acids

;; TYPE: amino acid

;; STRANDEDNESS: single

;; TOPOLOGY: linear

;; MOLECULE TYPE: No. 6232456e

US-08-944-483-62

Query Match 50.5%; Score 616.5; DB 4; Length 230;

Best Local Similarity 56.6%; Pred. No. 1.5e-52;

Matches 112; Conservative 31; Mismatches 54; Indels 1; Gaps 1;

QY 34 IYNGENAVLGSNPMQVSLQDSSGFHFCGSLISQSWVYTAHCNVSPGRHFVVLGEYDRS 93

DB 1 IYNGEDAVPGSNPMQVSLQDKTGFHFCGSLISEDWVYVTAHCGVRTS-DVVVAGEFDQ 59

QY 94 SNAEPLQVLSVSRATTHPSNMTNNNDVTLKLASPAQYTRISPCVCLASSNEALTEGL 153

DB 60 SDEENIQVLKTAKEFKNFPSILTVNNDITLLKATPARFQTSVAVCLPADDFFAGT 119

QY 154 TCVTGWRGLSGVGNVTPAHQQVALPLVTNQCQRYWDSITDSMICAGGAGSSCOG 213

DB 120 LCATTGWGKTKYANKTPDKLQQAALPLLSNAECKKSGRRITDVMICAGASGVSSCMGD 179

QY 214 DSGGPLVCQKGNWTWLVGI 231

DB 180 DSGGPLVCQKGNWTWLVGI 197

#### RESULT 4

US-08-557-146-13

;; Sequence 13, Application US/08557146

;; Patent No. 5834290

;; GENERAL INFORMATION:

;; APPLICANT: Egelrud, Torbjorn

;; APPLICANT: Hansson, Lennart

;; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic

;; TITLE OF INVENTION: Enzyme (SCCE)

;; NUMBER OF SEQUENCES: 17

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: White & Case, Patent Department

;; STREET: 1155 Avenue of the Americas

;; CITY: New York

;; STATE: New York

;; COUNTRY: U.S.A.

;; ZIP: 10036-2787

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/557,146

;; FILING DATE: 14-DEC-1995

;; CLASSIFICATION: 424

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Sterner, Richard J.

;; REGISTRATION NUMBER: 35,372

;; REFERENCE/DOCKET NUMBER: 1103326-181

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: (212) 819-8783

;; TELEFAX: (212) 354-8113



INFORMATION FOR SEQ ID NO: 13:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 229 amino acids  
 TYPE: amino acids  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: polypeptide  
 US-08-557-146-13

Query Match 49.9%; Score 609; DB 2; Length 229;  
 Best Local Similarity 57.1%; Pred. No. 8.2e-52;  
 Matches 113; Conservative 30; Mismatches 53; Indels 2; Gaps 2;  
 QY 34 IVNGENAVLGSWPQVSLQDSSGFHFCGSLISQSWVYTAACHNVSPPGRHFVWLGEYDRS 93  
 DB 1 IVNGEDAVPGSWPQVSLQDRTGPHFCGSLISEDWVYTAACHGVRTS-DVVVAGEFFDQG 59  
 QY 94 SNAEPQLVLSVSRATHTPSWNTMNDVTLLKLSPAQYTRISPVGLASSNEALTEGL 153  
 DB 60 SDEENIQVLKIAKVFKNPKFSILTVNNDITLLKLATPAREFQTVSAVCLPSADDDFPAGT 119  
 QY 154 TCVITGWRGLSGVGNVTPAHLQOVALPLVTVNQCQRYWDDSSITDSMICAGGAGASSCOGD 213  
 DB 120 LCATTGWGKTKYNANKTPDKLQQAALPLLSNAECKKSGRRITDYMICA-GAGVSSCMGD 178  
 QY 214 SGGPLVCOCKGNTWVLGI 231  
 DB 179 SGGPLVCOCKGAWTLVGI 196

RESULT 5  
 US-09-154-344-13  
 Sequence 13, Application US/09154344  
 Patent No. 5981256  
 GENERAL INFORMATION:  
 APPLICANT: Egelrud, Torbjorn  
 APPLICANT: Hansson, Lennart  
 TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic  
 TITLE OF INVENTION: Enzyme (SCCE)  
 NUMBER OF SEQUENCES: 17  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: White & Case, Patent Department  
 STREET: 1155 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: U.S.A.  
 ZIP: 10036-2787

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/154.344  
 FILING DATE: 16-SEP-1998  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/557,146  
 FILING DATE: 14-DEC-1995  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Sturner, Richard J.  
 REGISTRATION NUMBER: 35,372  
 REFERENCE/DOCKET NUMBER: 1103326-181  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 819-8783  
 TELEFAX: (212) 354-8113  
 INFORMATION FOR SEQ ID NO: 13:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 229 amino acids  
 TYPE: amino acids  
 STRANDEDNESS: single  
 TOPOLOGY: linear

MOLECULE TYPE: polypeptide  
 US-09-154-344-13

Query Match 49.9%; Score 609; DB 2; Length 229;  
 Best Local Similarity 57.1%; Pred. No. 8.2e-52;  
 Matches 113; Conservative 30; Mismatches 53; Indels 2; Gaps 2;  
 QY 34 IVNGENAVLGSWPQVSLQDSSGFHFCGSLISQSWVYTAACHNVSPPGRHFVWLGEYDRS 93  
 DB 1 IVNGEDAVPGSWPQVSLQDRTGPHFCGSLISEDWVYTAACHGVRTS-DVVVAGEFFDQG 59  
 QY 94 SNAEPQLVLSVSRATHTPSWNTMNDVTLLKLSPAQYTRISPVGLASSNEALTEGL 153  
 DB 60 SDEENIQVLKIAKVFKNPKFSILTVNNDITLLKLATPAREFQTVSAVCLPSADDDFPAGT 119  
 QY 154 TCVITGWRGLSGVGNVTPAHLQOVALPLVTVNQCQRYWDDSSITDSMICAGGAGASSCOGD 213  
 DB 120 LCATTGWGKTKYNANKTPDKLQQAALPLLSNAECKKSGRRITDYMICA-GAGVSSCMGD 178  
 QY 214 SGGPLVCOCKGNTWVLGI 231  
 DB 179 SGGPLVCOCKGAWTLVGI 196

RESULT 6  
 US-08-278-091-10  
 Sequence 10, Application US/08278091  
 Patent No. 5506139  
 GENERAL INFORMATION:  
 APPLICANT: LOOSMORE, Sheena M  
 APPLICANT: YANG, Yan-Ping  
 APPLICANT: CHONG, Pele  
 APPLICANT: OOMEN, Raymond P.  
 APPLICANT: KLEIN, Michel H.  
 TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with  
 TITLE OF INVENTION: Reduced Protease Activity  
 NUMBER OF SEQUENCES: 23  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sim & McBurney  
 STREET: Suite 701, 330 University Avenue  
 CITY: Toronto  
 STATE: Ontario  
 COUNTRY: Canada  
 ZIP: M5G 1R7

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/278.091  
 FILING DATE: 21-JUL-1994  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Stewart, Michael I  
 REGISTRATION NUMBER: 24,973  
 REFERENCE/DOCKET NUMBER: 1038-371  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (416) 595-1155  
 TELEFAX: (416) 595-1163  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 228 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear

US-08-278-091-10  
 Sequence 10, Application US/08278091  
 Patent No. 5506139  
 GENERAL INFORMATION:  
 APPLICANT: LOOSMORE, Sheena M  
 APPLICANT: YANG, Yan-Ping  
 APPLICANT: CHONG, Pele  
 APPLICANT: OOMEN, Raymond P.  
 APPLICANT: KLEIN, Michel H.  
 TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with  
 TITLE OF INVENTION: Reduced Protease Activity  
 NUMBER OF SEQUENCES: 23  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sim & McBurney  
 STREET: Suite 701, 330 University Avenue  
 CITY: Toronto  
 STATE: Ontario  
 COUNTRY: Canada  
 ZIP: M5G 1R7

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/278.091  
 FILING DATE: 21-JUL-1994  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Stewart, Michael I  
 REGISTRATION NUMBER: 24,973  
 REFERENCE/DOCKET NUMBER: 1038-371  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (416) 595-1155  
 TELEFAX: (416) 595-1163  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 228 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear

US-08-278-091-10  
 Sequence 10, Application US/08278091  
 Patent No. 5506139  
 GENERAL INFORMATION:  
 APPLICANT: LOOSMORE, Sheena M  
 APPLICANT: YANG, Yan-Ping  
 APPLICANT: CHONG, Pele  
 APPLICANT: OOMEN, Raymond P.  
 APPLICANT: KLEIN, Michel H.  
 TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with  
 TITLE OF INVENTION: Reduced Protease Activity  
 NUMBER OF SEQUENCES: 23  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sim & McBurney  
 STREET: Suite 701, 330 University Avenue  
 CITY: Toronto  
 STATE: Ontario  
 COUNTRY: Canada  
 ZIP: M5G 1R7

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/278.091  
 FILING DATE: 21-JUL-1994  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Stewart, Michael I  
 REGISTRATION NUMBER: 24,973  
 REFERENCE/DOCKET NUMBER: 1038-371  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (416) 595-1155  
 TELEFAX: (416) 595-1163  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 228 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear

US-08-278-091-10  
 Sequence 10, Application US/08278091  
 Patent No. 5506139  
 GENERAL INFORMATION:  
 APPLICANT: LOOSMORE, Sheena M  
 APPLICANT: YANG, Yan-Ping  
 APPLICANT: CHONG, Pele  
 APPLICANT: OOMEN, Raymond P.  
 APPLICANT: KLEIN, Michel H.  
 TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with  
 TITLE OF INVENTION: Reduced Protease Activity  
 NUMBER OF SEQUENCES: 23  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sim & McBurney  
 STREET: Suite 701, 330 University Avenue  
 CITY: Toronto  
 STATE: Ontario  
 COUNTRY: Canada  
 ZIP: M5G 1R7

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/278.091  
 FILING DATE: 21-JUL-1994  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Stewart, Michael I  
 REGISTRATION NUMBER: 24,973  
 REFERENCE/DOCKET NUMBER: 1038-371  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (416) 595-1155  
 TELEFAX: (416) 595-1163  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 228 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear

US-08-278-091-10  
 Sequence 10, Application US/08278091  
 Patent No. 5506139  
 GENERAL INFORMATION:  
 APPLICANT: LOOSMORE, Sheena M  
 APPLICANT: YANG, Yan-Ping  
 APPLICANT: CHONG, Pele  
 APPLICANT: OOMEN, Raymond P.  
 APPLICANT: KLEIN, Michel H.  
 TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with  
 TITLE OF INVENTION: Reduced Protease Activity  
 NUMBER OF SEQUENCES: 23  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sim & McBurney  
 STREET: Suite 701, 330 University Avenue  
 CITY: Toronto  
 STATE: Ontario  
 COUNTRY: Canada  
 ZIP: M5G 1R7

US-08-278-091-10  
 Sequence 10, Application US/08278091  
 Patent No. 5506139  
 GENERAL INFORMATION:  
 APPLICANT: LOOSMORE, Sheena M  
 APPLICANT: YANG, Yan-Ping  
 APPLICANT: CHONG, Pele  
 APPLICANT: OOMEN, Raymond P.  
 APPLICANT: KLEIN, Michel H.  
 TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with  
 TITLE OF INVENTION: Reduced Protease Activity  
 NUMBER OF SEQUENCES: 23  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sim & McBurney  
 STREET: Suite 701, 330 University Avenue  
 CITY: Toronto  
 STATE: Ontario  
 COUNTRY: Canada  
 ZIP: M5G 1R7

Db 1 IVNGEAVPGSWPQVSLQDKTGFHFCGSLINENWVVTAAHCGVTTSDVYVAGEFDQG 59  
QY 94 SNAEPLOVLSVRAITHPSWNTTMMNDVTLKSLASPAQYTTTRISPVCLASNEALTEGL 153  
Db 60 SSSEKIOKLIKIAKFNKSYNSLTINNDITLLKSLTAASFSQTVSAVCLPSASDDFAACT 119  
QY 154 TCVTTCGRLSGVGNVTPAHLOQVALPLVTVNOQRYWDSSTDSMICAGGAGASCQGD 213  
Db 120 TCVTTCGWL-LTRYAN-TPDLQOASLPLLSNTNCKKYNKTKIKDAMICAGASGVSSCMGD 177  
QY 214 SGGPLVCQKGNWTWVLGI 231  
Db 178 SGGPLVCKRNGAWTLVGI 195

## RESULT 7

US-08-483-859-10  
; Sequence 10, Application US/08483859  
; Patent No. 5656436  
; GENERAL INFORMATION:  
; APPLICANT: LOOSMORE, Sheena M.  
; APPLICANT: YANG, Yan-Ping  
; APPLICANT: CHONG, Pele  
; APPLICANT: COMEN, Raymond P.  
; APPLICANT: KLEIN, Michel H.  
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with  
; TITLE OF INVENTION: Reduced Protease Activity  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: Suite 701, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/483,859  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/296,149  
; FILING DATE: 26-AUG-1994

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/278,091  
; FILING DATE: 21-JUL-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I.  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-495 MIS:vg  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 228 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-483-859-10

Query Match 49.8%; Score 607.5; DB 1; Length 228;  
Best Local Similarity 55.6%; Pred. No. 1.1e-51;  
Matches 110; Conservative 36; Mismatches 49; Indels 3; Gaps 3;

QY 34 IVNGENAVLGSWPQVSLQDSSGFHFCGSLISQSWVVTAAHCNVSPGRHFVVLGEYDRS 93  
Db 1 IVNGEAVPGSWPQVSLQDKTGFHFCGSLINENWVVTAAHCGVTTSDVYVAGEFDQG 59

QY 94 SNAEPLOVLSVRAITHPSWNTTMMNDVTLKSLASPAQYTTTRISPVCLASNEALTEGL 153  
Db 60 SSSEKIOKLIKIAKFNKSYNSLTINNDITLLKSLTAASFSQTVSAVCLPSASDDFAACT 119  
QY 154 TCVTTCGRLSGVGNVTPAHLOQVALPLVTVNOQRYWDSSTDSMICAGGAGASCQGD 213  
Db 120 TCVTTCGWL-LTRYAN-TPDLQOASLPLLSNTNCKKYNKTKIKDAMICAGASGVSSCMGD 177  
QY 214 SGGPLVCQKGNWTWVLGI 231  
Db 178 SGGPLVCKRNGAWTLVGI 195

## RESULT 8

US-08-472-173-10  
; Sequence 10, Application US/08472173  
; Patent No. 5665353  
; GENERAL INFORMATION:  
; APPLICANT: LOOSMORE, Sheena M.  
; APPLICANT: YANG, Yan-Ping  
; APPLICANT: CHONG, Pele  
; APPLICANT: COMEN, Raymond P.  
; APPLICANT: KLEIN, Michel H.  
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with  
; TITLE OF INVENTION: Reduced Protease Activity  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: Suite 701, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/472,173  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/296,149  
; FILING DATE: 26-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/278,091  
; FILING DATE: 21-JUL-1994

ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I.  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-493 MIS:vg  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 228 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-472-173-10

Query Match 49.8%; Score 607.5; DB 1; Length 228;  
Best Local Similarity 55.6%; Pred. No. 1.1e-51;  
Matches 110; Conservative 36; Mismatches 49; Indels 3; Gaps 3;

QY 34 IVNGENAVLGSWPQVSLQDSSGFHFCGSLISQSWVVTAAHCNVSPGRHFVVLGEYDRS 93  
Db 1 IVNGEAVPGSWPQVSLQDKTGFHFCGSLINENWVVTAAHCGVTTSDVYVAGEFDQG 59  
QY 94 SNAEPLOVLSVRAITHPSWNTTMMNDVTLKSLASPAQYTTTRISPVCLASNEALTEGL 153

Db 60 SSSEKIQLKIAKVPKNSKYNLTNNITLLKLTAAASFQTVSAVCLPSASDDFAAGT 119  
Qy 154 TCVTGWLGRSLGSGVGNVTPAHLOQVALPLVTYNOCROYWDSSTDSMICAGAGASSCOGD 213  
Db 120 TCVTGWLGRSLGSGVGNVTPAHLOQVALPLVTYNOCROYWDSSTDSMICAGAGASSCOGD 177  
Qy 214 SGGPLVCKKNGNTWVLGI 231  
Db 178 SGGPLVCKKNGNTWVLGI 195

RESULT 9  
US-08-487-167-10  
; Sequence 10, Application US/08487167  
; Patent No. 5869302  
; GENERAL INFORMATION:  
; APPLICANT: LOOSMORE, Sheena M.  
; APPLICANT: YANG, Yan-Ping  
; APPLICANT: CHONG, Pele  
; APPLICANT: OOMEN, Raymond P.  
; APPLICANT: KLEIN, Michel H.  
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with  
; TITLE OF INVENTION: Reduced Protease Activity  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: Suite 701, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/487,167  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
; PRIOR APPLICATION NUMBER: US 08/296,149  
; FILING DATE: 26-AUG-1994  
; APPLICATION NUMBER: US 08/278,091  
; FILING DATE: 21-JUL-1994

ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I.  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-508 MIS:vg  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 228 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

US-08-487-167-10

Query Match 49.8%; Score 607.5; DB 2; Length 228;  
Best Local Similarity 55.6%; Pred. No. 1.1e-51;  
Matches 110; Conservative 36; Mismatches 49; Indels 3; Gaps 3;

Qy 34 IVNGENAVLGSWPMQVSLQDSGFHFCGSLISQSWVYTAHCNYSVGRHFVVLGEYDRS 93  
Db 1 IVNGEAVPGSWPMQVSLQDKTGFGHFCGSLINENWVYTAHCGVTTSDVVAAGEFDQG 59  
Qy 94 SNAEPQLVLSVSRATHTPSWNSNTMNDVTLKLTAAASFQTVSAVCLPSASDDFAAGT 153  
Db 60 SSSEKIQLKIAKVPKNSKYNLTNNITLLKLTAAASFQTVSAVCLPSASDDFAAGT 119

Qy 154 TCVTGWLGRSLGSGVGNVTPAHLOQVALPLVTYNOCROYWDSSTDSMICAGAGASSCOGD 213  
Db 120 TCVTGWLGRSLGSGVGNVTPAHLOQVALPLVTYNOCROYWDSSTDSMICAGAGASSCOGD 177  
Qy 214 SGGPLVCKKNGNTWVLGI 231  
Db 178 SGGPLVCKKNGNTWVLGI 195

RESULT 10

US-08-482-816-10  
; Sequence 10, Application US/08482816  
; Patent No. 5935573

GENERAL INFORMATION:  
; APPLICANT: LOOSMORE, Sheena M  
; APPLICANT: YANG, Yan-Ping  
; APPLICANT: CHONG, Pele  
; APPLICANT: OOMEN, Raymond P.  
; APPLICANT: KLEIN, Michel H.  
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with Reduced Protease A  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: Suite 701, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/482,816  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
; PRIOR APPLICATION NUMBER: US 08/296,149  
; FILING DATE: 26-AUG-1994  
; APPLICATION NUMBER: US 08/278,091  
; FILING DATE: 21-JUL-1994

ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-494 MIS:vg  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 228 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

US-08-482-816-10

Query Match 49.8%; Score 607.5; DB 2; Length 228;  
Best Local Similarity 55.6%; Pred. No. 1.1e-51;  
Matches 110; Conservative 36; Mismatches 49; Indels 3; Gaps 3;

Qy 34 IVNGENAVLGSWPMQVSLQDSGFHFCGSLISQSWVYTAHCNYSVGRHFVVLGEYDRS 93  
Db 1 IVNGEAVPGSWPMQVSLQDKTGFGHFCGSLINENWVYTAHCGVTTSDVVAAGEFDQG 59  
Qy 94 SNAEPQLVLSVSRATHTPSWNSNTMNDVTLKLTAAASFQTVSAVCLPSASDDFAAGT 153  
Db 60 SSSEKIQLKIAKVPKNSKYNLTNNITLLKLTAAASFQTVSAVCLPSASDDFAAGT 119  
Qy 154 TCVTGWLGRSLGSGVGNVTPAHLOQVALPLVTYNOCROYWDSSTDSMICAGAGASSCOGD 213  
Db 120 TCVTGWLGRSLGSGVGNVTPAHLOQVALPLVTYNOCROYWDSSTDSMICAGAGASSCOGD 177



US-08-615-271-10  
; Sequence 10, Application US/08615271  
; Patent No. 5981503  
; GENERAL INFORMATION:  
; APPLICANT: LOOSMORE, Sheena M.  
; APPLICANT: YANG, Yan-ping  
; APPLICANT: CHONG, Pele  
; APPLICANT: OOMEN, Raymond P.  
; APPLICANT: KLEIN, Michel H.  
; TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED  
; PROTEASE ACTIVITY  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: 6th Floor, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/615,271  
; FILING DATE: 20-JUN-1996  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I.  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-580  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 228 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-615-271-10

Query Match 49.8%; Score 607.5; DB 2; Length 228;  
Best Local Similarity 55.6%; Pred. No. 1.1e-51;  
Matches 110; Conservative 36; Mismatches 49; Indels 3; Gaps 3;  
Qy 34 IVNGENAVLGSWPQVSLQDSSGFHFCGSLISQSWVYTAACHCNVSPGRHFVILGEYDRS 93  
Db 1 IVNGEEAVPGSWPQVSLQDKTGFHFCGSLINENWVYTAACHCGVTS-DVVVAGEFDOG 59  
Qy 94 SNAEPQLVLSVSRATHPSWNSTNMNDVTLKLLASPAQYTRISPVCLASNEALTEGL 153  
Db 60 SSSEKIOKLIKAKFKNSKYNLSLTINDITLLKLTAAASFQTSVAVCLPSASDDFAAGT 119  
Qy 154 TCVTTGRLSGVGNVPAHLQVVALPLVTYNOCROYWDSITDSMICAGGAGASSCOGD 213  
Db 120 TCVTTGWL-LTRYAN-TPDRLOQASLPLLSNTNCKYKWTGKIKDAMICAGAGSVSSCMGD 177  
Qy 214 SGGPLVCOKGNTWVLIGI 231  
Db 178 SGGPLVCKKNGAWTLVGI 195

RESULT 14  
US-09-074-660-10  
; Sequence 10, Application US/09074660  
; Patent No. 6020183  
; GENERAL INFORMATION:  
; APPLICANT: LOOSMORE, Sheena M.  
; APPLICANT: YANG, Yan-ping  
; APPLICANT: CHONG, Pele  
; APPLICANT: OOMEN, Raymond P.

; APPLICANT: KLEIN, Michel H.  
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with  
; REDUCED Protease Activity  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: 6th Floor, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/074,660  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/487,167  
; FILING DATE: 07-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/296,149  
; FILING DATE: 26-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/278,091  
; FILING DATE: 21-JUL-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I.  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-731 MIS:jb  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 228 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-09-074-660-10

Query Match 49.8%; Score 607.5; DB 3; Length 228;  
Best Local Similarity 55.6%; Pred. No. 1.1e-51;  
Matches 110; Conservative 36; Mismatches 49; Indels 3; Gaps 3;  
Qy 34 IVNGENAVLGSWPQVSLQDSSGFHFCGSLISQSWVYTAACHCNVSPGRHFVILGEYDRS 93  
Db 1 IVNGEEAVPGSWPQVSLQDKTGFHFCGSLINENWVYTAACHCGVTS-DVVVAGEFDOG 59  
Qy 94 SNAEPQLVLSVSRATHPSWNSTNMNDVTLKLLASPAQYTRISPVCLASNEALTEGL 153  
Db 60 SSSEKIOKLIKAKFKNSKYNLSLTINDITLLKLTAAASFQTSVAVCLPSASDDFAAGT 119  
Qy 154 TCVTTGRLSGVGNVPAHLQVVALPLVTYNOCROYWDSITDSMICAGGAGASSCOGD 213  
Db 120 TCVTTGWL-LTRYAN-TPDRLOQASLPLLSNTNCKYKWTGKIKDAMICAGAGSVSSCMGD 177  
Qy 214 SGGPLVCOKGNTWVLIGI 231  
Db 178 SGGPLVCKKNGAWTLVGI 195

RESULT 15  
US-09-074-659-10  
; Sequence 10, Application US/09074659  
; Patent No. 6025342  
; GENERAL INFORMATION:  
; APPLICANT: LOOSMORE, Sheena M.  
; APPLICANT: YANG, Yan-ping  
; APPLICANT: CHONG, Pele

Search completed: December 20, 2002, 15:20:36  
Job time : 14.087 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:14:08 ; Search time 28.6087 Seconds  
(without alignments)  
1663.721 Million cell updates/sec

Title: US-09-856-319B-4\_COPY\_34\_264

Perfect score: 1252

Sequence: 1 IVNGENAVPGSWPQVSLQD.....AMVTRVSKFSTWVQVMAYN 231

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1252	100.0	264	11 Q9ER05	Q9er05 mus musculus
2	1248	99.7	264	11 Q9D7P8	Q9d7p8 mus musculus
3	1242	99.2	264	11 Q9D960	Q9d960 mus musculus
4	1199	95.8	264	11 Q9EQ28	Q9eq28 rattus norv
5	841.5	67.2	261	13 Q9W704	Q9w7q4 paralichthy
6	749.5	59.9	260	13 Q9W7Q3	Q9w7q3 paralichthy
7	727.5	58.1	263	11 Q9DC86	Q9dc86 mus musculus
8	725.5	57.9	263	11 Q9CR35	Q9cr35 mus musculus
9	719.5	57.5	263	11 Q9D8X8	Q9d8x8 mus musculus
10	712.5	56.9	263	13 Q9PWQ6	Q9pwq6 gadus morhu
11	674	53.8	164	11 Q9DC82	Q9dc82 mus musculus
12	511.5	40.9	269	11 Q9D7F9	Q9d7f9 mus musculus
13	511.5	40.9	269	11 Q9CQ52	Q9cq52 mus musculus
14	496	39.6	266	11 Q92077	Q92077 gadus morhu
15	494.5	39.5	269	6 Q95KW7	Q95kw7 bos taurus
16	493	39.4	249	13 Q9W7Q1	Q9w7q1 paralichthy

17	487	38.9	1524	13	Q91674	Q91674 xenopus lae
18	486.5	38.9	260	13	Q9W7P9	Q9w7p9 paralichthy
19	485.5	38.8	267	5	Q9BK47	Q9bk47 luidia foli
20	465.5	37.2	810	4	Q15146	Q15146 homo sapien
21	465	37.1	812	11	Q91WJ5	Q91wj5 mus musculu
22	462.5	36.9	266	13	Q9W7Q0	Q9w7q0 paralichthy
23	460	36.7	268	13	Q9W7Q2	Q9w7q2 paralichthy
24	459.5	36.7	270	4	Q96QL8	Q96ql8 homo sapien
25	457.5	36.5	257	6	O19023	O19023 macaca mula
26	457.5	36.5	277	5	O96899	O96899 scolopendra
27	453	36.2	269	4	Q96QV5	Q96qv5 homo sapien
28	451.5	36.1	339	11	Q99L44	Q99l44 mus musculu
29	448.5	35.8	273	11	Q921N4	Q921n4 mus musculu
30	444.5	35.5	334	6	O46507	O46507 papio hanad
31	444	35.5	812	11	Q9R0W3	Q9r0w3 rattus norv
32	442.5	35.3	270	13	Q91039	Q91039 gadus morhu
33	442.5	35.3	453	11	Q8VDR0	Q8vdr0 mus musculu
34	439.5	35.1	244	13	Q8QGW3	Q8qgw3 anguilla ja
35	438.5	35.0	266	6	O46644	O46644 macaca fasc
36	438.5	35.0	331	11	Q8RIA6	Q8ria6 mus musculu
37	438	35.0	246	11	Q9QUR9	Q9qur9 mus musculu
38	437.5	34.9	266	11	Q91X79	Q91x79 mus musculu
39	436.5	34.9	247	13	O42608	O42608 petromyzon
40	435.5	34.8	247	13	O42138	O42138 petromyzon
41	435.5	34.8	266	11	Q9D936	Q9d936 mus musculu
42	435	34.7	238	13	Q9W7Q6	Q9w7q6 paralichthy
43	435	34.7	246	11	Q921R9	Q921r9 mus musculu
44	434	34.7	246	11	Q9R0T7	Q9r0t7 mus musculu
45	432.5	34.5	806	6	O18783	O18783 macropus eu

#### ALIGNMENTS

#### RESULT 1

Q9ER05 ID Q9ER05 PRELIMINARY; PRT; 264 AA.  
AC Q9ER05;  
DT 01-MAR-2001 (TREMREL. 16, Created)  
DT 01-MAR-2001 (TREMREL. 16, Last sequence update)  
DT 01-JUN-2002 (TREMREL. 21, Last annotation update)  
DE Chymopasin (Chymotrypsin A CTRA-1).  
OS CTRL OR CTRL.  
GN Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Mitsui S., Yamaguchi N.;  
RT "Molecular cloning of mouse chymopasin."  
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129S6/SVEVTAC; TISSUE=SPLEEN;  
RA Bjoernslett M.;  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.  
DR EMBL; AB016228; BAB20275.1; -.  
DR EMBL; AF236365; AAL11034.1; -.  
DR HSSP; P00766; 4CHA.  
DR MEROPS; S01.256; -.  
DR MGD; MGI:88558; Ctrl.  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR001254; Ser\_protease\_Try.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease.

SQ SEQUENCE 264 AA; 28135 MW; 1D979709A07056C2 CRC64;  
 Query Match 100.0%; Score 1252; DB 11; Length 264;  
 Best Local Similarity 100.0%; Pred. No. 9.7e-107; Length 264;  
 Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IVNGENAVPGSPWQVSLQDNTGFEHFCGSLISPNVNVVTAACHQVTPGRHFVVLGEYDRS 60  
 DB 34 IVNGENAVPGSPWQVSLQDNTGFEHFCGSLISPNVNVVTAACHQVTPGRHFVVLGEYDRS 93  
 QY 61 SNAEPQVLSIARAIHPNNMNTMNDLTLLKLSPARYTAQVSPVCLASTNEALPSGL 120  
 DB 94 SNAEPQVLSIARAIHPNNMNTMNDLTLLKLSPARYTAQVSPVCLASTNEALPSGL 153  
 QY 121 TCVTGGRISGVNTPARLQOVLPVLTVNCROYWGARITDAMICAGGSCSQGD 180  
 DB 154 TCVTGGRISGVNTPARLQOVLPVLTVNCROYWGARITDAMICAGGSCSQGD 213  
 QY 181 SGGPLVCQKNTWVLIGIVSWGFKNCNIQAPAMYTRVSKFSTWVNOVMAYN 231  
 DB 214 SGGPLVCQKNTWVLIGIVSWGFKNCNIQAPAMYTRVSKFSTWVNOVMAYN 264

RESULT 2  
 Q9D7P8 PRELIMINARY; PRT; 264 AA.  
 ID Q9D7P8  
 AC Q9D7P8  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE 1810004D15RIK protein.  
 GN CTRL OR 1810004D15RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=TONGUE;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa K., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momaberts P.,  
 RA Nardone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 TRYPSIN FAMILY.  
 DR EMBL; AK009019; BAB26029.1; -.  
 DR HSSP; P00766; 4CHA.  
 DR MEROPS; S01.256; -.  
 DR MGD; MGI:88558; Ctrl.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR001254; Ser.protease\_Try.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; Tryp\_Spc; 1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.

RESULT 3  
 Q9D960 PRELIMINARY; PRT; 264 AA.  
 ID Q9D960  
 AC Q9D960  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE 1810004D15RIK protein.  
 GN CTRL OR 1810004D15RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=PANCREAS;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa K., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momaberts P.,  
 RA Nardone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 TRYPSIN FAMILY.  
 DR EMBL; AK007333; BAB24967.1; -.  
 DR HSSP; P00766; 4CHA.  
 DR MEROPS; S01.256; -.  
 DR MGD; MGI:88558; Ctrl.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR001254; Ser.protease\_Try.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; Tryp\_Spc; 1.



```

DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 264 AA; 28151 MW; 1D979719E07C16DE CRC64;

Query Match
Best Local Similarity 99.2%; Score 1242; DB 11; Length 264;
Matches 230; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVYTAHCKVTPGRHFVILGEYDRS 60
Db 34 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVYTAHCKVTPGRHFVILGEYDRS 93

Qy 61 SNAEPVQVLSARAIHPNNWNTMNDLTLKLSAPRYTAQVSPVCLASTNEALPSGL 120
Db 94 SNAEPVQVLSARAIHPNNWNTMNDLTLKLSAPRYTAQVSPVCLASTNEALPSGL 153

Qy 121 TCVTGGRISGVGNVTTPARLQVVLPLVTYNQCRQYWGARITDAMICAGSGASSCGD 180
Db 154 TCVTGGRISGVGNVTTPARLQVVLPLVTYNQCRQYWGARITDAMICAGSGASSCGD 213

Qy 181 SGGPLVCOKGNTWVLIGIVSGTKNCNIQAPAMYTRVSKFSTWYNQVAYN 231
Db 214 SGGPLVCOKGNTWVLIGIVSGTKNCNIQAPAMYTRVSKFSTWYNQVAYN 264

RESULT 4
Q9EQZ8 PRELIMINARY; PRT; 264 AA.
AC Q9EQZ8;
AT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chymopasin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-RAT PANCREAS;
RA Sogame Y., Mitsui S., Kataoka K., Kashima K., Kato M., Sakagami J.,
RA Yamaguchi N.;
RT "Molecular cloning of rat chymopasin."
RT Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AB020757; BAB20287.1; -.
DR HSSP; P00766; 4CHA.
DR MEROPS; S01.256; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 264 AA; 28116 MW; F9ED5D210FD3500E CRC64;

Query Match
Best Local Similarity 95.8%; Score 1199; DB 11; Length 264;
Matches 214; Conservative 16; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVYTAHCKVTPGRHFVILGEYDRS 60
Db 34 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVYTAHCKVTPGRHFVILGEYDRS 93

Qy 61 SNAEPVQVLSARAIHPNNWNTMNDLTLKLSAPRYTAQVSPVCLASTNEALPSGL 120
Db 94 SNAEPVQVLSARAIHPNNWNTMNDLTLKLSAPRYTAQVSPVCLASTNEALPSGL 153

```

```

Qy 121 TCVTGGRISGVGNVTTPARLQVVLPLVTYNQCRQYWGARITDAMICAGSGASSCGD 180
Db 154 TCVTGGRISGVGNVTTPARLQVVLPLVTYNQCRQYWGSRITDSMICAGAGASSCGD 213

Qy 181 SGGPLVCOKGNTWVLIGIVSGTKNCNIQAPAMYTRVSKFSTWYNQVAYN 231
Db 214 SGGPLVCOKGNTWVLIGIVSGTKNCNIQAPAMYTRVSKFSTWYNQVAYN 264

RESULT 5
Q9W7Q4 PRELIMINARY; PRT; 261 AA.
AC Q9W7Q4;
AT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chymotrypsinogen 1.
OS Paralicthys olivaceus (Flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectidae; Paralicthidae; Paralicthys.
OX NCBI_TaxID=8255;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PANCREAS;
RA Suzuki T., Srivastava A.S., Kurokawa T.;
RT "Japanese flounder mRNA for chymotrypsinogen 1."
RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AB029753; BAA82365.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.256; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 261 AA; 28184 MW; D7090A9D65395B7D CRC64;

Query Match
Best Local Similarity 65.9%; Pred. No. 3.7e-69;
Matches 153; Conservative 28; Mismatches 48; Indels 3; Gaps 2;

Qy 1 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVYTAHCKVTPGRHFVILGEYDRS 60
Db 32 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVYTAHCKVTPGRHFVILGEYDRS 91

Qy 61 SNAEPVQVLSARAIHPNNWNTMNDLTLKLSAPRYTAQVSPVCLASTNEALPSGL 120
Db 92 YNNEPIQVMSARAIHPNNWNTMNDLTLKLSAPRYTAQVSPVCLASTNEALPSGL 151

Qy 121 TCVTGGRISGVGNVTTPARLQVVLPLVTYNQCRQYWG-ARITDAMICAGSGASSCG 179
Db 152 KCVTTGGRGTQTS--SPRYLQQTSLPLSPAQCKQYGNRITDAMICAGSGVSCQG 209

Qy 180 DSGGPLVCOKGNTWVLIGIVSGTKNCNIQAPAMYTRVSKFSTWYNQVAYN 231
Db 210 DSGGPLVCOKGNTWVLIGIVSGTKNCNIQAPAMYTRVSKFSTWYNQVAYN 261

RESULT 6
Q9W7Q3 PRELIMINARY; PRT; 260 AA.
AC Q9W7Q3;
AT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

```

DT 01-WAR-2002 (TREMELrel. 20, Last annotation update)  
DE Chymotrypsinogen 2.  
OS Paralicthys olivaceus (Flounder).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;  
OC Pleuronectoidae; Paralicthidae; Paralicthidae; Paralicthys.  
OX NCBI\_TaxID=8255;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PANCREAS;  
RA Suzuki T., Srivastava A.S., Kurokawa T.;  
RT "Japanese flounder mRNA for chymotrypsinogen 2.";  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
CC TRYPSIN FAMILY.  
DR EMBL; AB029754; BAB82366.1; --  
DR HSSP; P00766; 1CGT.  
DR MEROPS; S01.152; --  
DR InterPro; IPR001314; Chymotrypsin.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; trypsin; 1.  
DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease.  
SQ SEQUENCE 260 AA; 27793 MW; 9F583044E22F78C0 CRC64;  
Query Match 59.9%; Score 749.5; DB 13; Length 260;  
Best Local Similarity 57.6%; Pred. No. 9.8e-61;  
Matches 133; Conservative 35; Mismatches 62; Indels 1; Gaps 1;  
QY 1 IVNGENAVPGSPWQVSLQDNTGFHFCGSLSPNVTAAHCQVTPGRHFVVLGEYDRS 60  
Db 31 IVNGEALPHSPWQVSLQDNTGFHFCGSLSPNVTAAHCNVRTS-HRVILGEHDS 89  
QY 61 SNAEPQVLSIARAIHPNNMANTMNDLTLKLPASARYTAQVSPVCLASTNEALPSGL 120  
Db 90 SNAEDIQVVKVGVKPRYNGTYTINDLLIKLAAPAMNHRVSPVCVAETSDNFAGM 149  
QY 121 TCVTGTGWRISGVNTPARLQVVLPLVTNQCQRYGARTDAMICAGSGSCQGD 180  
Db 150 KCVTSGWGLTRNAPDPTPALQQAALPLTNDCCRYGKISNLICAGSAGSCSCMD 209  
QY 181 SGGPLVCQKGNVTWLVIGVSWGTKNICQAPAMYTRVSKFSTWVNOYMAYN 231  
Db 210 SGGPLVCQKAGATLVIGVSWGSGTCTPTMPGYARVTELRAWMDQTIANN 260  
RESULT 7  
Q9DC86 PRELIMINARY; PRT; 263 AA.  
ID Q9DC86;  
AC Q9DC86;  
DT 01-JUN-2001 (TREMELrel. 17, Created)  
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)  
DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)  
DE 220008D09rik protein.  
GN 220008D09rik.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=SPLEEN;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Aizawa K., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boileau D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
CC TRYPSIN FAMILY.  
DR EMBL; AK003060; BAB22539.1; --  
DR HSSP; P00766; 1GCT.  
DR MEROPS; S01.152; --  
DR MGD; MGI:1913723; 220008D09rik.  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR001254; Ser\_protease\_Try.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; Tryp\_Spc; 1.  
DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease.  
SQ SEQUENCE 263 AA; 27821 MW; 2620A27AFBA5D04D CRC64;  
Query Match 58.1%; Score 727.5; DB 11; Length 263;  
Best Local Similarity 56.3%; Pred. No. 1e-58;  
Matches 130; Conservative 37; Mismatches 63; Indels 1; Gaps 1;  
QY 1 IVNGENAVPGSPWQVSLQDNTGFHFCGSLSPNVTAAHCQVTPGRHFVVLGEYDRS 60  
Db 34 IVNGEDAIPGSPWQVSLQDRTGFHFCGSLSPNVTAAHCQVKT-TNVVAGEFDG 92  
QY 61 SNAEPQVLSIARAIHPNNMANTMNDLTLKLPASARYTAQVSPVCLASTNEALPSGL 120  
Db 93 SDEENQVQLKIAQVKNPKFNFTVRNDITLKLATPAQFSETSAVCLFTVDDDFPAGT 152  
QY 121 TCVTGTGWRISGVNTPARLQVVLPLVTNQCQRYGARTDAMICAGSGSCQGD 180  
Db 153 LCATTGKGTKNYALKTPDKLQQAALPIVSEAKCKESGSKITDYMICAGSSVSCMD 212  
QY 181 SGGPLVCQKGNVTWLVIGVSWGTKNICQAPAMYTRVSKFSTWVNOYMAYN 231  
Db 213 SGGPLVCQKDGWTLGIVSWGSGFCSTSTPAVYARVATLMPWQVLEIAN 263  
RESULT 8  
Q9CR35 PRELIMINARY; PRT; 263 AA.  
ID Q9CR35;  
AC Q9CR35;  
DT 01-JUN-2001 (TREMELrel. 17, Created)  
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)  
DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)  
DE 220008D09rik protein.  
GN 220008D09rik.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=STOMACH, SPLEEN, AND PANCREAS;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX	MEDLINE=21085660; PubMed=11217851;
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA	Saito T., Okazaki Y., Gotojori T., Bono H., Kasukawa T., Saito R.,
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA	Fleischmann W., Gaasterland T., Glissi C., King B., Kochiwa H.,
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA	Gustinch S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA	L Lyons P., Marchionni L., Mashima J., Mazzarelli J., Wombaerts P.,
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA	Zsuzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
RA	Hayashizaki Y.;
RT	"Functional annotation of a full-length mouse cDNA collection.";
RL	Nature 409:685-690(2001)
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.
DR	EMBL; AK007566; BAB25112.1; -;
DR	HSSP; P00766; LGCT.
DR	MEROPS; S01.152; -;
DR	MGD; MGI:1913723; Z20008BD09Rik.
DR	InterPro; IPRO01314; Chymotrypsin.
DR	InterPro; IPRO01254; Ser_protease_Try.
DR	Pfam; PF00089; trypsin; 1.
DR	PRINTS; PR00722; CHYMOTRYPSIN.
DR	SMART; SM00020; Tryp_SPC; 1.
DR	PROSITE; PS02040; TRYPSIN_DOM; 1.
DR	PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR	PROSITE; PS00135; TRYPSIN_SER; 1.
KW	Hydrolase. Serine protease.
SQ	SEQUENCE 263 AA; 27898 MW; C0638FBF905A92F CRC64;
Query Match                57.5%; Score 719.5; DB 11; Length 263;	
Best Local Similarity 55.8%; Pred No. 5.5e-58;	
Matches 129; Conservative 36; Mismatches 68; Indels 1; Gaps	
Oy	1 IVNGENAVPGSWPMQVSLDNTGFHFCGGLSPNNKVVTAAHCHQVTGPRHFYVLGEYDRS 60
Dd	34 IVNGEIDAIPGSWMQVSLDRTGFHFCGGYLIVSNVNVTAAHCGVKYT-TDVVVAGEFDQG 92
Oy	61 SNAEPVOVLSTARATHPNKNATMNDLTLLKLASPARYTQAQVPVCLASTNEALPSGL 120
Dd	93 SDEENVQLAKIAQVFKNPKFSFTVRNDITLULKLATPAQSETVSACVLTPTVDDFFPAT 152
Oy	121 TCVTGTWGRIISGVGNVTPAPLKQWLPLVTVINQCROYGCARITDMICAGSGASSCOGD 180
Dd	153 LCATTGWGKTNYNALTKPDKLQAALLPIVESEAKRESWGSKITDWNICAGASGYSSCMGD 212
Oy	181 SGGPLVCQKGNTWLVIGIWSWGTKNCNIQAPAMTYTRVSKFTWIINVWAYN 231
Dd	213 SGGPLVCQKGVMTWLAGIWSWGGFGCSTSTPAVIARTVATMPWVBOILEAN 263

RESULT 10	
Q9PWQ6	
ID	Q9PWQ6
AC	PRELIMINARY; PRT; 263 AA.
DT	Q9PWQ6;
DT	01-MAY-2000 (TrEMBLrel. 13, Created)
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE	Chymotrypsin B precursor (EC 3.4.21.1).
GN	CHYB.
OS	Gadus morhua (Atlantic cod).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthopterygii; Paracanthopterygii; Gadiformes; Gadidae; Gadus.  
 OX NCBI\_TaxID=8049;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PYLORIC CAECA;  
 RA MEDLINE=20464334; PubMed=11011764;  
 RA Spilliaert R., Gudmundsdottir A.;  
 RT "Molecular Cloning of the Atlantic Cod Chymotrypsinogen B";  
 RL Microb. Comp. Genomics 5:41-50(2000).  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPSIN FAMILY.  
 DR EMBL; AJ242521; CAB43766.1; -.  
 DR HSP; P00766; ICHG.  
 DR MEROPS; S01.152; -.  
 DR InterPro: IPR001314; Chymotrypsin.  
 DR InterPro: IPR001254; Ser\_protease\_Try.  
 DR Pfam; PF00089; trypsin; 1.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS0134; TRYPSIN\_HIS; UNKNOWN\_1.  
 DR PROSITE; PS0135; TRYPSIN\_SER; 1.  
 KW Hydrolase; Serine protease; Signal.  
 FT SIGNAL 1 16  
 FT CHAIN 17 263 CHYMOTRYPSIN B.  
 FT SEQUENCE 263 AA; 28175 MW; EF61B18A34EE57C CRC64;  
 SQ  
 Query Match 56.98; Score 712.5; DB 13; Length 263;  
 Best Local Similarity 54.9%; Pred. No. 2.4e-57;  
 Matches 129; Conservative 41; Mismatches 58; Indels 7; Gaps 5;  
 QY 1 IVNGENAVPGSWQVSLQDNTGTFHFGGSLISPNVYTAACHCOVTPGRHFWVLGEYDRS 60  
 DB 32 IVNGEAVPHSWQVSLQDSNGFHCGLINENWVYTAACHVRT-YHRVIIVEGHDKS 90  
 QY 61 -SNAEPVOVLSTARAITHPNNTNNTMNDLTLKLPARYTAQVSPVCLASTNEALPSG 119  
 DB 91 RASDNIQILKPSMVFTHPKWDSTINNDISLIKLPASVLTGTVNVPCLGESSDVFAPG 150  
 QY 120 LTCVTTGG--RISGVGNVTPARLQOVPLVLTYNOCROYWGR-ITDAMICAGSGASS 176  
 DB 151 MKCVTSGLGRYRNAP--TPNKIQOALPLMSNECSQTWGNMISDMVICAGAAGATS 208  
 QY 177 CQGDGGLVCGKQNTWTLVIGVSWGTNCNIOAPAMYTRVSKFTSWINQVMAYN 231  
 DB 209 CMGDSGGLVCGKQNTWTLVIGVSWGSRCSVTTPAYARYTELGRWVDQILAN 263  
 RESULT 11  
 ID Q9DC82 PRELIMINARY; PRT; 164 AA.  
 AC Q9DC82;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE 0910001G08Rik protein.  
 GN CTRL OR 0910001G08RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=SPLEEN;  
 RA MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.;  
 RA "Functional annotation of a full-length mouse cDNA collection";  
 RL Nature 409:685-690(2001).  
 DR EMBL; AK003074; BAB22549.1; -.  
 DR HSP; P00763; IDPO.  
 DR MEROPS; S01.997; -.  
 DR MGD; MGI:88558; Ctrl.  
 DR InterPro: IPR001254; Ser\_protease\_Try.  
 DR Pfam; PF00089; trypsin; 1.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS0135; TRYPSIN\_SER; 1.  
 KW Hydrolase; Serine protease.  
 SQ SEQUENCE 164 AA; 17707 MW; 83791FD929ABEDD6 CRC64;  
 Query Match 53.8%; Score 674; DB 11; Length 164;  
 Best Local Similarity 82.2%; Pred. No. 4.5e-54;  
 Matches 129; Conservative 5; Mismatches 21; Indels 2; Gaps 2;  
 QY 77 HPNNA-NTWNNDLTLKLPAS-ARYTAQVSPVCLASTNEALPSGLTCVTGWRISGVG 134  
 DB 8 HPWHPGLKTKTRLFLEFFPHPKTKFPFGLASTNEALPSGLTCVTGWRISGVG 67  
 QY 135 NVTPLQOVPLVLTYNOCROYWGRITDAMICAGSGASSCGDSGLVCGKQNTW 194  
 DB 68 NVTPLQOVPLVLTYNOCROYWGRITDAMICAGSGASSCGDSGLVCGKQNTW 127  
 QY 195 LIGVSWGTNCNIOAPAMYTRVSKFTSWINQVMAYN 231  
 DB 128 LIGVSWGTNCNIOAPAMYTRVSKFTSWINQVMAYN 164  
 RESULT 12  
 ID Q9D7T9 PRELIMINARY; PRT; 269 AA.  
 AC Q9D7T9;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE 2310074F01Rik protein.  
 GN ELA3B OR 2310074F01RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=STOMACH;  
 RA MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

[illegible]

RX MEDLINE-96262679; PubMed-8925447;  
RA Gudmundsdottir E., Spilliaert R., Yang Q., Craik C.S., Bjarnason J.B.,  
RA "Isolation and characterization of two cDNAs from Atlantic cod  
RT encoding two distinct psychrophilic elastases";  
RL Comp. Biochem. Physiol. 113B:795-801(1996).  
CC -I- CATALYTIC ACTIVITY: HYDROLYSIS OF PROTEINS, INCLUDING ELASTIN.  
CC PREFERENTIAL CLEAVAGE: VAL-I-XAA > ALA-I-XAA.  
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
CC TRYPSIN FAMILY.  
CC EMBL: U56936; AAB38350.1; -.  
CC HSP: P05805; IFON.  
DR MEROPS: S01.155; -.  
DR InterPro: IPR001314; Chymotrypsin.  
DR InterPro: IPR001254; Ser.protease\_Try.  
DR Pfam: PF00089; trypsin; 1.  
DR PRINTS: PR00722; CHYMOTRYPSIN.  
DR SMART: SM00020; TRYPSIN\_DOM; 1.  
DR PROSITE: PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE: PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease; Signal; zymogen.  
FT SIGNAL 1 15 POTENTIAL.  
FT PROPEP 16 27 ACTIVATION PEPTIDE (POTENTIAL).  
FT CHAIN 28 266 ELASTASE.  
FT CHAIN 28 266 ELASTASE.  
SQ SEQUENCE 266 AA; 28533 MW; B786B52C71559E2E CRC64;  
  
Query Match 39.6%; Score 496; DB 13; Length 266;  
Best Local Similarity 42.4%; Pred. No. 1.6e-37;  
Matches 101; Conservative 48; Mismatches 77; Indels 12; Gaps 9;  
  
QY 1 IVNGENAVPCSWPQVSLDNTG---FHFCGSLISPNWVTAACHQVTPGRHFVVLGEY 57  
DB 28 VVGEDVRVHSWFQASLQKSNYSYHCGGLIAPQWVWTAHC-IGSRTYRVLLGKH 86  
  
QY 58 D-RSSNAEPVQVLSIARAIHPNNANTMNDTLKLSAPARYTAQVSPVCLASTNEAL 116  
DB 87 NMQDYNEAGSLAIPAKIIVHEKWDSSRIKDIALIKLASPDVDSAIITPACVPDAEVL 146  
  
QY 117 PSLGLTCVTTGWGRISGVNVTAPLQOVPLPLVTVNCQRY--WGARITDAMICAGSGA 174  
DB 147 ANGAPCYVTGWRLTGGPIADA-LQALLPVVDHACRSYDWMGSLVTTSMYCAGGDGV 205  
  
QY 175 -SSCGDGGPLVCQKGN-TWVLIGIVSWGCK-NCNI-QAPAMYTRVSKFSTWVNOVM 228  
DB 206 LASCNGDGGPLNCQADGSDWVHGVSFGSSMGVCNPKPSVTRVSAYIPINNVN 263  
  
RESULT 15  
Q95KW7 PRELIMINARY; PRT; 269 AA.  
AC Q95KW7;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Proproteinase E (Fragment).  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PANCREAS;  
RA Seddi R., Guo X.-J., Chaix J.-C., Puigserver A.;  
RT "Nucleotide sequence of a bovine pancreatic proproteinase E cDNA."  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY057840; AAL23697.1; -.  
DR InterPro: IPR001254; Ser.protease\_Try.  
DR Pfam: PF00089; trypsin; 1.  
DR PROSITE: PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE: PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
DR PROSITE: PS00135; TRYPSIN\_SER; UNKNOWN\_1.

KW Hydrolase; Serine protease.  
FT NON\_TER 1  
SQ SEQUENCE 269 AA; 29010 MW; 352E4202B04B8B2D CRC64;  
  
Query Match 39.5%; Score 494.5; DB 6; Length 269;  
Best Local Similarity 43.2%; Pred. No. 2.3e-37;  
Matches 105; Conservative 37; Mismatches 88; Indels 13; Gaps 8;  
  
QY 1 IVNGENAVPCSWPQVSLQ---DNTGFHFCGSLISPNWVTAACHQVTPGRHFVVLGEY 57  
DB 28 VNGEDAVPYSWQVSLQYKDGAFHHTCGGSLIAPDWVVTAGHCILSTRTYQVVLGEY 87  
  
QY 58 DRS--SNAEPVQVLSIARAIHPNNANTM--NNDTLKLSAPARYTAQVSPVCLASTN 113  
DB 88 DRSVLESGEQVIPINAGDLFVHPLWNSNCVACGNDIALYKLSRSQAQGLDKVOLANLPPAG 147  
  
QY 114 EALPSGLTCVTTGWGRISGVNVTAPLQOVPLPLVTVNCQRY--WGARITDAMICAGG 171  
DB 148 DILPNEAPCYISGGRLY-TGGPLPKLQKALLPVVDYEHCSQWDMWGITVKKTMVCAGG 206  
  
QY 172 SGASSCGDGGPLVCQKGN-TWVLIGIVSW-GTKNCN-IQAPAMYTRVSKFSTWVNOVM 228  
DB 207 DTRSGCNGDGGPLNCPADGSMQVHGVTSFVSFGCNTIKKPTVTRVSFIDWIDETI 266  
  
QY 229 AYN 231  
DB 267 ASN 269

Search completed: December 20, 2002, 15:18:57  
Job time : 29.6087 secs

Result 3

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:12:13 ; Search time 35.6087 Seconds  
(without alignments)  
864.421 Million cell updates/sec

Title: US-09-856-319B-2\_COPY\_1\_231

Perfect score: 1221  
Sequence: 1 MLLSLTSLVLLGSSWGG.....GDSGGPLVCQKGNWTWVLI 231

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*  
1: /SID22/gcgdata/geneseq/geneseq-embl/AA1980.DAT.\*  
2: /SID22/gcgdata/geneseq/geneseq-embl/AA1981.DAT.\*  
3: /SID22/gcgdata/geneseq/geneseq-embl/AA1982.DAT.\*  
4: /SID22/gcgdata/geneseq/geneseq-embl/AA1983.DAT.\*  
5: /SID22/gcgdata/geneseq/geneseq-embl/AA1984.DAT.\*  
6: /SID22/gcgdata/geneseq/geneseq-embl/AA1985.DAT.\*  
7: /SID22/gcgdata/geneseq/geneseq-embl/AA1986.DAT.\*  
8: /SID22/gcgdata/geneseq/geneseq-embl/AA1987.DAT.\*  
9: /SID22/gcgdata/geneseq/geneseq-embl/AA1988.DAT.\*  
10: /SID22/gcgdata/geneseq/geneseq-embl/AA1989.DAT.\*  
11: /SID22/gcgdata/geneseq/geneseq-embl/AA1990.DAT.\*  
12: /SID22/gcgdata/geneseq/geneseq-embl/AA1991.DAT.\*  
13: /SID22/gcgdata/geneseq/geneseq-embl/AA1992.DAT.\*  
14: /SID22/gcgdata/geneseq/geneseq-embl/AA1993.DAT.\*  
15: /SID22/gcgdata/geneseq/geneseq-embl/AA1994.DAT.\*  
16: /SID22/gcgdata/geneseq/geneseq-embl/AA1995.DAT.\*  
17: /SID22/gcgdata/geneseq/geneseq-embl/AA1996.DAT.\*  
18: /SID22/gcgdata/geneseq/geneseq-embl/AA1997.DAT.\*  
19: /SID22/gcgdata/geneseq/geneseq-embl/AA1998.DAT.\*  
20: /SID22/gcgdata/geneseq/geneseq-embl/AA1999.DAT.\*  
21: /SID22/gcgdata/geneseq/geneseq-embl/AA2000.DAT.\*  
22: /SID22/gcgdata/geneseq/geneseq-embl/AA2001.DAT.\*  
23: /SID22/gcgdata/geneseq/geneseq-embl/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1221	100.0	264	21 AAB11710	Human serine prote
2	1093	89.5	264	21 AAB11711	Mouse serine prote
3	996	81.6	192	21 AAB34077	Human pancreatic c
4	684.5	56.1	263	23 AAU82738	Amino acid sequenc
5	657.5	53.8	245	21 AAY99596	Bovine chymotrypsi
6	621.5	50.9	231	22 AAB98504	Human chymotrypsin
7	472	38.7	269	7 AAP61724	Porcine elastase I
8	472	38.7	269	8 AAP70758	Pig pancreas elast
9	446.5	36.6	253	13 AAR29621	Porcine pancreatic
10	445	36.4	269	7 AAP60062	Sequence Of human

11	439.5	36.0	269	7 AAP61723	Human elastase II.
12	427.5	35.0	269	21 AAB54124	Human pancreatic c
13	427	35.0	269	8 AAP70760	Human pancreas ela
14	427	35.0	1052	22 AAM78338	Human protein seq
15	424	34.7	319	21 AAB36481	Fusion gene with h
16	424	34.7	319	22 AAB67541	Anino acid sequenc
17	422	34.6	252	7 AAP60058	Sequence of human
18	420.5	34.4	270	8 AAP70759	Human pancreas ela
19	420.5	34.4	270	23 AAU87691	Human pancreatic t
20	420.5	34.4	273	21 AAB54243	Human pancreatic c
21	418.5	34.3	253	7 AAP60059	Sequence of human
22	417	34.2	279	22 ABG20513	Novel human diagno
23	416.5	34.1	242	7 AAP60061	Sequence of human
24	414.5	33.9	269	7 AAP61076	Human elastase I.
25	410.5	33.6	343	23 AAU78547	/ Human prostatic pr
26	410.5	33.6	343	23 ABB07285	Anino acid sequenc
27	410.5	33.6	453	21 AAB44250	Human PRO382 (UNQ3
28	408.5	33.5	452	20 AAY41694	Human PRO382 (prote
29	408.5	33.5	453	22 AAU29055	Human PRO polypept
30	408.5	33.5	453	22 AAE06935	Human membrane-ty
31	408.5	33.5	453	23 AAE33020	Human trypsin fami
32	408.5	33.5	454	21 AAB32246	Tumour associated
33	408.5	33.5	454	22 AAU68911	Human TAGD-12, pro
34	408	33.4	248	21 AAB43572	Human cancer assoc
35	408	33.4	327	21 AAY72093	Human serine prote
36	408	33.4	327	23 AAE17921	Human gene 3 encod
37	408	33.4	394	23 ABB41994	Human ovarian anti
38	408	33.4	454	23 AAE33024	Human trypsin fami
39	408	33.4	454	23 AAU82745	Anino acid sequenc
40	407.5	33.4	268	17 AAR90682	Rat caldesrin cont
41	403.5	33.0	811	12 AAR12933	Plasminogen mutain
42	403.5	33.0	811	12 AAR12945	Plasminogen mutain
43	401.5	32.9	811	12 AAR12939	Plasminogen mutain
44	401.5	32.9	811	12 AAR12943	Plasminogen mutain
45	399.5	32.7	811	12 AAR12944	Plasminogen mutain

## ALIGNMENTS

### RESULT 1

AAB11710  
ID AAB11710 standard; Protein; 264 AA.

XX AAB11710;

XX 23-OCT-2000 (first entry)

DE Human serine protease BSSP5 (hbSSP5) SEQ ID NO:2.

XX BSSP5: serine protease; human; hbSSP5; mouse; mBSSP5; brain;  
KW diagnostic marker; antibody; transgenic animal; Alzheimer's disease;  
KW epilepsy; cancer; inflammation; infertility; pancreatitis;  
KW prostatic hypertrophy.

OS Homo sapiens.

PN WO200031243-A1.

XX 02-JUN-2000.

XX 19-NOV-1999; 99WO-JP06473.

XX 20-NOV-1998; 98JP-0347806.

XX (FUSO ) FUSO PHARM IND LTD.

PI Uemura H, Okui A, Kominami K, Yamaguchi N, Mitsui S;

XX WPI: 2000-400058/34.

DR N-PSDB; AAA61733.

XX Serine proteases BSSP5, useful in detecting homologs, mutants and



PT polymorphic variants as markers for diagnosis of e.g. Alzheimer's  
PT disease, epilepsy, cancer and inflammation, using blood, urine,  
PS pancreas or other tissues -  
XX  
XX Claim 1; Page 51-52; 70pp; Japanese.  
XX  
CC The invention relates to novel serine proteases designated BSSP5  
CC (AAB11710-B11711), and to nucleic acids encoding them (AAA61733-A61734).  
CC The invention also relates to vectors and transformatants comprising BSSP5  
CC nucleic acids; transgenic animals in which the expression level of BSSP5  
CC can be varied; and an mBSSP5 knockout mouse. The invention additionally  
CC encompasses anti-BSSP5 antibodies and methods of production of such  
CC antibodies, methods of BSSP5 detection using the antibodies, and the  
CC use of BSSP5 proteins or fragments as diagnostic markers for certain  
CC medical conditions, e.g., pancreatitis. A method for detecting  
CC pancreatitis comprising measuring BSSP5 concentration in the blood or  
CC urine, and a pancreatitis diagnostic agent containing an anti-BSSP5  
CC antibody is also disclosed. Nucleotides encoding BSSP5 were initially  
CC isolated in a human brain cDNA library using degenerate PCR primers  
CC (AAA61744-A61745) based on conserved regions of serine proteases. The  
CC BSSP5 serine proteases and nucleotides encoding them are useful in  
CC detecting homologues, mutants and polymorphic variants in biological  
CC samples (e.g. blood, urine, brain, prostate gland, placenta, testis,  
CC pancreas and spleen) as diagnostic markers for conditions such as  
CC Alzheimer's disease, epilepsy, cancer, inflammation, infertility,  
CC pancreatitis and prostatic hypertrophy. Sequence AAB11710 represents  
CC human BSSP5 (hBSSP5), and sequence AAB11711 represents murine BSSP5  
CC (mBSSP5).  
XX  
XX Sequence 264 AA;  
XX  
XX Query Match 100.0%; Score 1221; DB 21; Length 264;  
XX Best Local Similarity 100.0%; Pred. No. 1.1e-98;  
XX Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLLSLTSLVLGSGGCGIPAIKPAISQRIVNGENAVLGSWPQVSLQDSSGFHFC 60  
DB 1 MLLSLTSLVLGSGGCGIPAIKPAISQRIVNGENAVLGSWPQVSLQDSSGFHFC 60  
QY 61 GGSLSQSWVVTAAHCNVSPGRHFVVLGVEYDRSSNAEPQLQVLSVRAITHPSWNSTMMN 120  
DB 61 GGSLSQSWVVTAAHCNVSPGRHFVVLGVEYDRSSNAEPQLQVLSVRAITHPSWNSTMMN 120  
QY 121 DVTLLKLPAAQVTTTRISPVCLASSNEALTEGLTCVTTGWRGSLGVGNVTPAHLQOVALP 180  
DB 121 DVTLLKLPAAQVTTTRISPVCLASSNEALTEGLTCVTTGWRGSLGVGNVTPAHLQOVALP 180  
QY 181 LVTVNOCROYWDSITDSMTICAGGAGASSCGDGGPLVCQKNTWVLIGI 231  
DB 181 LVTVNOCROYWDSITDSMTICAGGAGASSCGDGGPLVCQKNTWVLIGI 231  
RESULT 2  
AAB11711  
ID AAB11711 standard; Protein; 264 AA.  
XX  
XX AAB11711;  
XX  
XX 23-OCT-2000 (first entry)  
XX  
XX Mouse serine protease BSSP5 (mBSSP5) SEQ ID NO:4.  
XX  
XX BSSP5; serine protease; human; hBSSP5; mouse; mBSSP5; brain;  
XX diagnostic marker; antibody; transgenic animal; Alzheimer's disease;  
XX epilepsy; cancer; inflammation; infertility; pancreatitis;  
XX prostatic hypertrophy.  
XX  
XX Mus sp.  
XX  
XX WO200031243-A1.  
XX  
XX 02-JUN-2000.  
XX

PF 19-NOV-1999; 99WO-JP06473.  
XX  
PR 20-NOV-1998; 98JP-0347806.  
XX  
PA (FUSO ) FUSO PHARM IND LTD.  
XX  
PI Uemura H, Okui A, Kominami K, Yamaguchi N, Mitsui S;  
XX  
XX WPI; 2000-400058/34.  
DR N-PSDB; AAA61734.  
DR  
XX Serine proteases BSSP5, useful in detecting homologs, mutants and  
PT polymorphic variants as markers for diagnosis of e.g. Alzheimer's  
PT disease, epilepsy, cancer and inflammation, using blood, urine,  
PT pancreas or other tissues -  
XX  
XX Claim 3; Page 55-56; 70pp; Japanese.  
XX  
XX The invention relates to novel serine proteases designated BSSP5  
CC (AAB11710-B11711), and to nucleic acids encoding them (AAA61733-A61734).  
CC The invention also relates to vectors and transformatants comprising BSSP5  
CC nucleic acids; transgenic animals in which the expression level of BSSP5  
CC can be varied; and an mBSSP5 knockout mouse. The invention additionally  
CC encompasses anti-BSSP5 antibodies and methods of production of such  
CC antibodies, methods of BSSP5 detection using the antibodies, and the  
CC use of BSSP5 proteins or fragments as diagnostic markers for certain  
CC medical conditions, e.g., pancreatitis. A method for detecting  
CC pancreatitis comprising measuring BSSP5 concentration in the blood or  
CC urine, and a pancreatitis diagnostic agent containing an anti-BSSP5  
CC antibody is also disclosed. Nucleotides encoding BSSP5 were initially  
CC isolated in a human brain cDNA library using degenerate PCR primers  
CC (AAA61744-A61745) based on conserved regions of serine proteases. The  
CC BSSP5 serine proteases and nucleotides encoding them are useful in  
CC detecting homologues, mutants and polymorphic variants in biological  
CC samples (e.g. blood, urine, brain, prostate gland, placenta, testis,  
CC pancreas and spleen) as diagnostic markers for conditions such as  
CC Alzheimer's disease, epilepsy, cancer, inflammation, infertility,  
CC pancreatitis and prostatic hypertrophy. Sequence AAB11710 represents  
CC human BSSP5 (hBSSP5), and sequence AAB11711 represents murine BSSP5  
CC (mBSSP5).  
XX  
XX Sequence 264 AA;  
XX  
XX Query Match 89.5%; Score 1093; DB 21; Length 264;  
XX Best Local Similarity 85.7%; Pred. No. 1.6e-87;  
XX Matches 198; Conservative 21; Mismatches 12; Indels 0; Gaps 0;  
QY 1 MLLSLTSLVLGSGGCGIPAIKPAISQRIVNGENAVLGSWPQVSLQDSSGFHFC 60  
DB 1 MLLSLTSLVLGSGGCGIPAITPALTSTNQRIVNGENAVPGSWPQVSLQDTEGFHFC 60  
QY 61 GGSLSQSWVVTAAHCNVSPGRHFVVLGVEYDRSSNAEPQLQVLSVRAITHPSWNSTMMN 120  
DB 61 GGSLSIPNWWVTAAHCNVTPGRHFVVLGVEYDRSSNAEPQVLSIARAITHPNNANTMMN 120  
QY 121 DVTLLKLPAAQVTTTRISPVCLASSNEALTEGLTCVTTGWRGSLGVGNVTPAHLQOVALP 180  
DB 121 DLTLLKLPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQOVALP 180  
QY 181 LVTVNOCROYWDSITDSMTICAGGAGASSCGDGGPLVCQKNTWVLIGI 231  
DB 181 LVTVNOCROYGARIITDAMICAGSGASSCGDGGPLVCQKNTWVLIGI 231  
RESULT 3  
AAB54077  
ID AAB54077 standard; Protein; 192 AA.  
XX  
XX AAB54077;  
XX  
XX 09-MAR-2001 (first entry)  
XX  
XX Human pancreatic cancer antigen protein sequence SEQ ID NO:529.  
XX



XX Human; pancreas; pancreatic cancer; pancreatic cancer antigen;  
KW detection; diagnosis; identification; cytostatic; neuroprotective;  
KW neoplastic; immunomodulatory; relaxant; contraceptive; gynaecological;  
KW antiinflammatory; cardiant; gene therapy; chromosome mapping;  
KW linkage analysis; tissue identification; tissue typing; forensic;  
KW neural; immune system; muscular; reproductive; gastrointestinal;  
KW pulmonary; cardiovascular; renal; proliferative.  
XX  
OS Homo sapiens.  
XX WO200055320-A1.  
XX 21-SEP-2000.  
XX  
XX 08-MAR-2000; 2000WO-US05989.  
XX  
XX 12-MAR-1999; 99US-0124270.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Ruben SM;  
XX WPI: 2000-579444/54.  
XX N-PSDB; AAC98842.  
XX  
XX New nucleic acid that is a pancreatic cancer antigen for preventing,  
PT treating, or ameliorating a medical condition, particular pancreatic  
PT cancer, or for use in assays for diagnosing a pathological condition -  
XX  
XX Claim 11; Page 966; 1379pp; English.  
XX  
XX AAC98773 to AAC99231 encode the human pancreatic cancer associated  
CC proteins, called pancreatic cancer antigens, given in AAB54008 to  
CC AAB54466. The human pancreatic cancer antigens have cytostatic,  
CC neuroprotective, nootropic, immunomodulatory, relaxant, contraceptive,  
CC gynaecological, cardiant and antiinflammatory activities, and can be used  
CC in gene therapy. The polynucleotide and proteins can be used for  
CC preventing, treating, or ameliorating a medical condition or in assays  
CC for diagnosing a pathological condition or a susceptibility to one in a  
CC subject. Binding partners to the proteins and the activity of the  
CC proteins can be identified. The pancreatic cancer antigens can be used to  
CC detect, treat or prevent pancreatic disorders, especially cancer.  
CC Agonists and antagonists to the antigens can be screened for. The  
CC pancreatic cancer antigen polynucleotides can be used to design nucleic  
CC acid hybridisation probes that can be used in chromosome mapping, linkage  
CC analysis, tissue identification and/or typing and a variety of forensic  
CC and diagnostic methods. The proteins can be used to generate antibodies  
CC which are used to purify, detect and target the polypeptides, including  
CC both in vivo and in vitro diagnostic and therapeutic methods. The  
CC proteins can be used to treat or prevent neural, immune system, muscular,  
CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or  
CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent  
CC sequences used in the exemplification of the present invention.  
XX  
XX Sequence 192 AA;  
SQ  
Query Match 81.6%; Score 996; DB 21; Length 192;  
Best Local Similarity 99.5%; Pred. No. 3.3e-79;  
Matches 189; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 5 SUTLSLVLLGSSWGGCIPAIKPAKLSFSQRIVNGENAVLGSPWQVSLQDSSGFHFCGSL 64  
DB 1 SUTLSLVLLGSSWGGCIPAIKPAKLSFSQRIVNGENAVLGSPWQVSLQDSSGFHFCGSL 60  
QY 65 ISQSWVVTAAHCNVSGRHFVVLGEGYDRSSNAEPLOVLSVSRATHPSWNSSTMMNDVTL 124  
DB 61 ISQSWVVTAAHCNVSGRHFVVLGEGYDRSSNAEPLOVLSVSRATHPSWNSSTMMNDVTL 120  
QY 125 LKIASPAQVYTRISPVCLASSNEALTEGLTCVTTGNGRLSGVGNVTPAHLQVVALPLTV 184  
DB 121 LKIASPAQVYTRISPVCLASSNEALTEGLTCVTTGNGRLSGVGNVTPAHLQVVALPLTV 180

QY 185 NCRQYWDSS 194  
DB 181 NCRQYWGSS 190  
RESULT 4  
AAU82738  
ID AAU82738 standard; Protein; 263 AA.  
XX AAU82738;  
XX 23-APR-2002 (first entry)  
XX Amino acid sequence of novel human protease #37.  
DE Human; protease; cancer; immune-related disorder; cardiovascular disease;  
XX neuronal-associated disease; metabolic disorder; inflammatory disorder;  
KW nervous system disorder; sexual dysfunction; pain; mood disorder;  
KW hypertension; psychotic disorder; neurological disorder; dyskinesia;  
KW viral infection; human immunodeficiency virus; HIV; non-viral infection;  
KW ocular disease; cytostatic; enzyme.  
XX  
XX Homo sapiens.  
XX WO200200860-A2.  
XX 03-JAN-2002.  
XX 26-JUN-2001; 2001WO-US20171.  
XX 26-JUN-2000; 2000US-214047P.  
XX (SUGE-) SUGEN INC.  
XX Plowman G, Whyte D, Sudarsanam S, Manning G, Caenepeel S;  
PI Charyczak G;  
XX WPI: 2002-139913/18.  
XX N-PSDB; ABK31780.  
XX Nucleic acids encoding novel human proteases, useful for useful for  
PT treating diseases and disorders such as cancers, immune-related  
PT diseases and disorders, cardiovascular disease (e.g. restenosis) and  
PT inflammatory disorders -  
XX  
XX Claim 6; Fig 2M; 313pp; English.  
XX The present invention relates to the isolation of novel human  
CC proteases, and the nucleic acids encoding them. The sequences of  
CC the invention are useful for treating diseases and disorders such as  
CC cancers (e.g. breast, colon, lung), immune-related diseases and disorders  
CC (e.g. inflammatory diseases and asthma), cardiovascular diseases  
CC (e.g. restenosis and coronary thrombosis), brain or neuronal-associated  
CC diseases, metabolic disorders (e.g. diabetes, obesity), inflammatory  
CC disorders (e.g. rheumatoid arthritis and psoriasis), central or  
CC peripheral nervous system diseases, migraines, pain, sexual dysfunction,  
CC mood disorders, attention disorders, cognition disorders, hypotension,  
CC hypertension, psychotic disorders, neurological disorders  
CC (e.g. Alzheimer's disease, Parkinson's disease) and dyskinesias.  
CC The nucleic acids and polypeptides are also useful for treating viral  
CC infections caused by human immunodeficiency virus (HIV), and non-viral  
CC infections such as ocular disease (e.g. glaucoma) and macular  
CC degeneration. AAU82702-AAU82760 represent the novel human proteases of  
CC the invention.  
XX  
XX Sequence 263 AA;  
SQ  
Query Match 56.1%; Score 684.5; DB 23; Length 263;  
Best Local Similarity 55.0%; Pred. No. 8e-52;  
Matches 127; Conservative 36; Mismatches 67; Indels 1; Gaps 1;  
QY 1 MLLLSLTLVLGSSWGGCIPAIKPAKLSFSQRIVNGENAVLGSPWQVSLQDSSGFHFC 60

Db 1 MASLWLSGSLVGAARCGGVPATHPVLVSLSRVINGEDAVPGSWPQVSLQDKTGFHFC 60  
 QY 61 GGSLSOSVWYTAARCNVSPGRHVPVLGEYDRSSNAEPLOVLSVSRATHPSWNTMNN 120  
 Db 61 GGSLSISDWWTAARHCGVRRS-DVVVAGEFDQSGDENIQVLKIAKFKPKPSILTVNN 119  
 QY 121 DVTLLKIASPAQYTRISPCVCLASSNEALTEGLTCVTITGWRSLGSGVGNVTPAHLLQOVALP 180  
 Db 120 DITLLKATPARFQSVAVCLPSADDDFAGTLCAATGKTKYNANKTPDKLQQAALP 179  
 QY 181 LVTYNOCROYWDSITSMICAGAGASSCGQSDSGGLVPCOKGNTWVLIGI 231  
 Db 180 LLSNAECKSMGRRITDVMICAGAGSVSSCMGSDGGLVPCOKDGAWTLVGI 230  
 RESULT 5  
 AAY99596  
 ID AAY99596 standard; protein; 245 AA.  
 AC AAY99596;  
 DT 13-SEP-2000 (first entry)  
 DE Bovine chymotrypsinogen A.  
 XX Bovine; plasminogen activator; cardiant; thrombolytic;  
 KW heart attack; stroke; blood clotting disorder.  
 XX Bos taurus.  
 XX W0200032759-A1.  
 XX 08-JUN-2000.  
 PD 06-MAY-1999; 99WO-US09991.  
 PF 02-DEC-1998; 98US-0110588.  
 PR (OKLA-) OKLAHOMA MEDICAL RES FOUND.  
 XX Lin X, Zhang XC, Tang JUN;  
 PI WPI; 2000-422975/36.  
 DR Polyptide with plasminogen activator activity useful as thrombolytic  
 XX agent for treating blood clot disorders e.g. heart attack, comprises 10  
 PT amino acid peptide fragment for recognition or activation of  
 PT plasminogen  
 XX Disclosure; Page 35-36; 41pp; English.  
 XX The present sequence is bovine chymotrypsinogen. It was included in  
 CC a review of sequence homologies of several plasminogen activators.  
 CC Plasminogen is the principal serine protease zymogen in the  
 CC extracellular fluids of vertebrates. Its active form, plasmin, is  
 CC implicated in pericellular proteolysis associated with a wide range of  
 CC physiological and pathological processes. Plasminogen expression is  
 CC regulated by plasminogen activators which hydrolyse a peptide bond in  
 CC plasminogen to convert it to plasmin or form tight binding complexes  
 CC with plasminogen to spontaneously convert it to plasmin. The sequence  
 CC homology analysis has identified a six amino acid peptide involved in  
 CC plasminogen activation. This peptide is particularly useful when inserted  
 CC between amino acid residues 644 and 645 of full length human  
 CC plasminogen. Novel plasminogen activators have been made based upon the  
 CC plasminogen activation/recognition site of plasminogen binding proteins.  
 CC The polypeptides are useful in preparing thrombolytic agents for treating  
 CC blood clotting disorders such as heart attack.  
 XX Sequence 245 AA;

Query Match 53.8%; Score 657.5; DB 21; Length 245;  
 Best Local Similarity 54.9%; Pred. No. 1.7e-49;  
 Matches 117; Conservative 37; Mismatches 58; Indels 1; Gaps 1;

QY 19 CGIPAIPALSFQRIVNGENAVLGSWPQVSLQDSSGFHFCGSLISQSWSVTAACHNV 78  
 Db 1 CGVPAIQVLSGLSRVINGEAVPGSWPQVSLQDKTGFHFCGSLINENWVYTAACHGV 60  
 QY 79 SPGRHFVVLGEYDRSSNAEPLOVLSVSRATHPSWNTMNNVTLKLASPAQYTRIS 138  
 Db 61 TTS-DVVVAGEFDQSGSSERIKLAKVFPKSKYNSLTINNDITLLKLSAASFQTS 119  
 QY 139 PVLCLASSNEALTEGLTCVTITGWRSLGSGVGNVTPAHLLQOVALPVTYNOCROYWDSITDS 198  
 Db 120 AVLCPASDDFAAGTTCVTITGWRSLTRYTNANTPRLQOASLPLLSNTNCKRYWGKTKRDA 179  
 QY 199 MICAGAGASSCGQSDSGGLVPCOKGNTWVLIGI 231  
 Db 180 MICAGAGSVSSCMGSDGGLVPCOKGNTWVLIGI 212  
 RESULT 6  
 AAB98504  
 ID AAB98504 standard; Protein; 231 AA.  
 AC AAB98504;  
 DT 03-AUG-2001 (first entry)  
 DE Human chymotrypsin serine protease catalytic domain.  
 XX Human; TADG-15; cytostatic; vaccine; ovarian tumour; cancer;  
 KW tumour antigen-derived gene 15; serine protease; chymotrypsin.  
 XX Homo sapiens.  
 XX W0200129056-A1.  
 PD 26-APR-2001.  
 PF 20-OCT-2000; 2000WO-US29095.  
 PR 20-OCT-1999; 99US-0421213.  
 XX (UYAR-) UNIV ARKANSAS.  
 XX O'Brien TJ, Tanimoto H;  
 PI WPI; 2001-381031/40.  
 DR Novel extracellular serine protease, termed tumor antigen-derived gene  
 PT 15 protein overexpressed in carcinomas and DNA encoding it, for  
 PT diagnosis, treatment, prevention of cancer, particularly breast,  
 PT ovarian cancer  
 XX Example 10; Fig 1; 130pp; English.  
 XX The present invention relates to human tumour antigen-derived gene 15  
 CC (TADG-15) protein and coding sequence (see AAH23601 and AAB98500).  
 CC TADG-15 is an extracellular serine protease. It was found that TADG-15 is  
 CC over-expressed in ovarian tumours. TADG-15 protein or its fragments of  
 CC 9-20 residues that lack TADG-15 protease activity are useful for  
 CC vaccinating an individual against TADG-15, having, suspected of having or  
 CC at risk of getting cancer. Furthermore, the TADG-15 gene can be used as a  
 CC diagnostic or therapeutic target in cancer. The present sequence was used  
 CC in a sequence homology alignment with the catalytic domain of TADG-15.  
 XX Sequence 231 AA;

Query Match 50.9%; Score 621.5; DB 22; Length 231;  
 Best Local Similarity 56.8%; Pred. No. 2.2e-46;  
 Matches 113; Conservative 31; Mismatches 54; Indels 1; Gaps 1;

QY 33 RYVNGENAVLGSWPQVSLQDSSGFHFCGSLISQSWSVTAACHNVSPGRHFVVLGEYDR 92  
 Db 1 RYVNGENAVLGSWPQVSLQDSSGFHFCGSLISQSWVTAACHNVSPGRHFVVLGEYDR 59

Qy 93 SSNAPLQVLSVRAITHPSWNTTMMNDVITLLKLSAPQYTRISPVCLASSNEALTEG 152  
 Db 60 GSDEENIQVLAKVFKPKPSILTVNDITLLKLTAPRSQTVSAVCLPSADDFPAG 119  
 Qy 153 LTCVTTGNGRLSGVGNVTPAHLLQVALPLVTVNQCRQYWDSSITDSMICAGAGASSCOG 212  
 Db 120 TLCAITGNGKTKYNANKTPDKLQAALPLLSNAECKSKWGRRTITDVMICAGASGVSCMG 179  
 Qy 213 DSGGPLVCOKGNWVVLIGI 231  
 Db 180 DSGGPLVCOKDGAWTLVGI 198

RESULT 7  
 AAP61724  
 ID AAP61724 standard; Protein: 269 AA.  
 XX  
 AC AAP61724;  
 XX  
 DT 23-OCT-1991 (first entry)  
 XX  
 DE Porcine elastase II.  
 XX  
 KW Colibacillus; yeast.  
 XX  
 OS Sus scrofa.  
 XX  
 PN JP61192289-A.  
 XX  
 PD 26-AUG-1986.  
 XX  
 PF 22-FEB-1985; 85JP-0034050.  
 XX  
 PR 22-FEB-1985; 85JP-0034050.  
 XX  
 PA (KIRI ) KIRIN BREWERY KK.  
 XX  
 WPI: 1986-262895/40.  
 DR  
 N-PSDB; AAN60919.  
 XX  
 PT Biological prepn. of pig elastase II - using DNA with base  
 PT sequence coding for specified aminoacid sequence.  
 XX  
 PS Claim 1; Fig 1; 11pp; Japanese.  
 CC  
 CC The porcine elastase product may be efficiently expressed from a  
 CC transformed host such as colibacillus or yeast, yielding the product  
 CC more efficiently than spleen extraction.  
 CC See also J61192288.  
 XX  
 SQ Sequence 269 AA;

Query Match 38.7%; Score 472; DB 7; Length 269;  
 Best Local Similarity 45.6%; Pred. No. 3e-33;  
 Matches 109; Conservative 32; Mismatches 84; Indels 14; Gaps 9;  
 Qy 2 LLLSLTSLVLGSSWGGCIPAIKPAISFQRIVNGENAVLGSWPQVSLQ--DSSG--FH 58  
 Db 1 MIRALLLTSLVAG--ALSCGLPANLPQ---PRVVGEDARPNSWPQVSLQYDSSGQWRH 56  
 Qy 59 FCGGSLISQSWVYTAHCNVSPGRHFVVLGEYDRSSNAEPQLQVLSVRAITHPSWNTTM 118  
 Db 57 TCGGTLVDQSWLTAHCNISSTRTYRVVGLRHSLSSTNEPGSLAVKLVVHQDNNSQL 116  
 Qy 119 N--NDVTLLKLSAPQYTRISPVCLASSNEALTEGLTCVTTGWGRLSGVGNVTPAHLLQ 176  
 Db 117 SKGNDIALKLSAPVLTDKIQLGCLPAAGSLPNNYVCYVTGWGRLQTNG--ASPDILQ 175  
 Qy 177 VALPLVTYNQCRQ--YWDSSITDSMICAGAG--ASSCGDGGPLVCOKGN--TWVLIGI 231  
 Db 176 GOLLVVDYATCSKPGWGSTVTKTNMICAGDGIISCSNGDSGGLNCGCQAGQWQVHGI 234

RESULT 8  
 AAP70758  
 ID AAP70758 standard; Protein: 269 AA.  
 XX  
 AC AAP70758;  
 XX  
 DT 19-APR-1991 (first entry)  
 XX  
 DE Pig pancreas elastase-2.  
 XX  
 KW Lipoprotein metabolism; lipid.  
 XX  
 OS Sus scrofa.  
 FH Key Location/Qualifiers  
 FT Misc-difference 2..16  
 FT /note= "May be absent"  
 XX  
 PN JP62000276-A.  
 PD 06-JAN-1987.  
 XX  
 PF 25-JUN-1985; 85JP-0138494.  
 XX  
 PR 25-JUN-1985; 85JP-0138494.  
 XX  
 PA (SANY ) SANKYO KK.  
 XX  
 WPI: 1987-040875/06.  
 DR  
 N-PSDB; AAN71121.  
 XX  
 PT Pancreas elastase prodn. used to improve lipoprotein metabolism -  
 PT comprises isolating RNA coding elastase, synthesising single and  
 PT double chain cDNA and introducing recombinant into host  
 XX  
 PS Disclosure; Page 480; 18pp; Japanese.  
 CC  
 CC Sequence encoding pancreas elastase may be used to transform an  
 CC expression system. The product is useful in rejuvenating the  
 CC elasticity of the arterial wall, and improving unusual serum lipid  
 CC levels and lipoprotein metabolism.  
 XX  
 SQ Sequence 269 AA;  
 Query Match 38.7%; Score 472; DB 8; Length 269;  
 Best Local Similarity 45.6%; Pred. No. 3e-33;  
 Matches 109; Conservative 32; Mismatches 84; Indels 14; Gaps 9;  
 Qy 2 LLLSLTSLVLGSSWGGCIPAIKPAISFQRIVNGENAVLGSWPQVSLQ--DSSG--FH 58  
 Db 1 MIRALLLTSLVAG--ALSCGLPANLPQ---PRVVGEDARPNSWPQVSLQYDSSGQWRH 56  
 Qy 59 FCGGSLISQSWVYTAHCNVSPGRHFVVLGEYDRSSNAEPQLQVLSVRAITHPSWNTTM 118  
 Db 57 TCGGTLVDQSWLTAHCNISSTRTYRVVGLRHSLSSTNEPGSLAVKLVVHQDNNSQL 116  
 Qy 119 N--NDVTLLKLSAPQYTRISPVCLASSNEALTEGLTCVTTGWGRLSGVGNVTPAHLLQ 176  
 Db 117 SKGNDIALKLSAPVLTDKIQLGCLPAAGSLPNNYVCYVTGWGRLQTNG--ASPDILQ 175  
 Qy 177 VALPLVTYNQCRQ--YWDSSITDSMICAGAG--ASSCGDGGPLVCOKGN--TWVLIGI 231  
 Db 176 GOLLVVDYATCSKPGWGSTVTKTNMICAGDGIISCSNGDSGGLNCGCQAGQWQVHGI 234

RESULT 9  
 AAR29621  
 ID AAR29621 standard; Protein: 253 AA.  
 XX  
 AC AAR29621;  
 XX  
 DT 30-APR-1993 (first entry)

```

XX DE Porcine pancreatic elastase III.
XX KW Pig; swine; bile acid; gall bladder; bile acid secretion promoters;
XX KW liver function improvers; ss.
XX OS Sus scrofa domestica.
XX PN JP04325090-A.
XX PD 13-NOV-1992.
XX PF 23-APR-1991; 91JP-0092069.
XX PR 23-APR-1991; 91JP-0092069.
XX PA (SANY ) SANKYO CO LTD.
XX WPI; 1992-428826/52.
DR N-PSDB; AAQ31724.
XX PT Pig pancreas elastase protein - used in bile acid secretion
XX PT promoters and liver function improvers
XX PS Claim 1; Page 8; 10pp; Japanese.
XX CC This sequence represents porcine pancreatic elastase III. It was
XX CC expressed in E. coli YA21 using the expression vector pELF001.
XX CC It may be used in bile acid secretion promoters and liver function
XX CC improvers.
XX SQ Sequence 253 AA;

Query Match 36.6%; Score 446.5; DB 13; Length 253;
Best Local Similarity 43.5%; Pred. No. 4.8e-31;
Matches 97; Conservative 33; Mismatches 78; Indels 15; Gaps 8;

QY 19 CGIPAIKPAKLSFSQIRIVNGENAVLGSMPQVSLQ-DSSGF--HFCGGLISQSWVYTAH 75
DB 1 CGRPSYNPA----SRVNGEDAVPYSWPQVSLQYEKNGVFQHTCGGLIAPDWVLTAGH 56
QY 76 CNVSPGRHFVVLGEYDRSSNAEPLQVLSYRA--ITHPSWNTMN--NDVTLKILASPA 131
DB 57 C1SSSTYQVVLGEYDRSENEGFEQVPIPNAGDLFVHPRWNSNCVSGNDIALVILKLSRA 116
QY 132 QYTRISPVCLASSNEALTEGLTCVTGWRGLSGVGNVTPAHLQQVALPLVTYNQCRQ-- 189
DB 117 QLGDKVQLACLPPAGDILPNDTPCYISGWRGLYTNGLP-PDKLQALLPVVDYQHCWKD 175
QY 190 YWDSSTIDSMICAGGAGASSCGDGGPLVCQKGN-TWVLIGI 231
DB 176 WVGSTVKQTWVCAGGDIRSNGDSGGLNCPAAGSGWQVHGV 218

RESULT 10
AAP60062
ID AAP60062 standard; Protein; 269 AA.
XX AC AAP60062;
XX AC AAP60062;
XX DT 06-SEP-1991 (first entry)
XX DE Sequence of human pancreatic elastase IIA encoded on pH2E2.
XX KW Enzyme; serum lipoprotein metabolism.
XX OS Homo sapiens.
XX PN EP198645-A.
XX PD 22-OCT-1986.
XX PF 07-APR-1986; 86EP-0302557.

```

```

XX 23-OCT-1985; 85JP-0236686.
PR 05-APR-1985; 85JP-0072308.
PR 27-APR-1985; 85JP-0091986.
PR 26-JUL-1985; 85JP-0163964.
PR 02-DEC-1985; 85JP-0271128.
XX (SANY ) SANKYO KK.
XX Takiguchi Y, Tani T, Kawashima I, Eurukawa H, Ohmine T;
PI Ohsumi J;
XX WPI; 1986-280300/43.
DR N-PSDB; AAN60076.
XX New genetically engineered human pancreatic elastase - obtcd.
PT using hosts modified DNA coding for enzyme
XX Example; Page 14-15; 45pp; English.
XX The genetically engineered prod. can eliminate the dependency on
CC human pancreas supplies for the elastase, and avoids antibody
CC formation and possibility of anaphylaxis using porcine elastase.
XX SQ Sequence 269 AA;

Query Match 36.4%; Score 445; DB 7; Length 269;
Best Local Similarity 43.9%; Pred. No. 6.9e-31;
Matches 105; Conservative 28; Mismatches 92; Indels 14; Gaps 8;

QY 2 LLSLTLSLVLLGSSWCGGIPAIKPAKLSFSQIRIVNGENAVLGSMPQVSLQDSSG---FH 58
DB 1 MIRTLLSTIVAG-ALSCGDPTYP---YVTVYVGGEARPNPMPQVSLQYSSNGKWH 56
QY 59 FCGGLISQSWVYTAACHCNVSPGRHFVVLGEYDRSSNAEPLQVLSYRAITHPSWNTM 118
DB 57 TCGGLISQSWVYTAACHCNVSPGRHFVVLGEYDRSSNAEPLQVLSYRAITHPSWNTM 116
QY 119 N--NDVTLKILASPAQYTRISPVCLASSNEALTEGLTCVTGWRGLSGVGNVTPAHLQQ 176
DB 117 SKGNDIALKLPVSLTDKIQIACLPAGDILPNDTPCYISGWRGLYTNGLP-PDKLQALLPVVDYQHCWKD 175
QY 177 VALPLVTYNQCRQ--YWDSSITDSMICAGGAGASSCGDGGPLVCQKGN-TWVLIGI 231
DB 176 GRLLVVDYATCSSAWWSSVKTSMICAGDGVISSNGDSGGLNCPAAGSGWQVHGI 234

RESULT 11
AAP61723
ID AAP61723 standard; Protein; 269 AA.
XX AC AAP61723;
XX DT 23-OCT-1991 (first entry)
XX DE Human elastase II.
XX KW Colibacillus; yeast.
XX OS Homo sapiens.
XX PN JP61192288-A.
XX PD 26-AUG-1986.
XX PF 22-FEB-1985; 85JP-0034049.
XX PR 22-FEB-1985; 85JP-0034049.
XX (KIRI ) KIRIN BREWERY KK.
XX WPI; 1986-262894/40.
DR N-PSDB; AAN60706.

```

PT New nucleic acid that is a pancreatic cancer antigen for preventing,  
PT treating, or ameliorating a medical condition, particular pancreatic  
PT cancer, or for use in assays for diagnosing a pathological condition

XX

cancer, or for use in assays for diagnosing a pathological condition.



```

XX 30-APR-1999: 99US-0303162.
XX
XX (ORTH ) ORTHO-MCNEIL PHARM RES INC.
XX
XX Darrow A, Qi J, Andrade-Gordon P;
XX
XX WPI; 2000-687533/67.
XX
XX Expression vector for producing recombinantly producing serine protease
XX domains, comprising a prosequence, a prosequence, and a cloning site
XX for the insertion of catalytic domain cassette -
XX
XX Disclosure: Page 47-50; 89pp; English.
XX
XX The present invention describes an expression vector (I) comprising in
XX frame and in order, a prosequence, a prosequence, and a cloning site
XX for the in frame insertion of catalytic domain cassette. (I) can be
XX used as a modulator of proteins expressed from a zymogen activation
XX construct. The recombinant catalytic domain of serine protease is
XX useful for identifying compounds modulating the activity of proteases
XX is expressed and activated from the zymogen activation construct. A
XX method from the present invention comprises combining a modulator of
XX the recombinant catalytic domain of a protease and measuring an effect
XX of the modulator on the protein preferably inhibiting or enhancing its
XX enzymatic activity or stimulation or inhibition of proteolysis mediated
XX by the expressed catalytic domain. The present sequence represents a
XX fusion gene with a human serine protease catalytic domain protein
XX sequence from the present invention.
XX
SQ Sequence 319 AA;
Query Match 34.7%; Score 424; DB 21; Length 319;
Best Local Similarity 35.9%; Pred. No. 5.8e-29;
Matches 92; Conservative 39; Mismatches 97; Indels 28; Gaps 7;
QY 1 MLLSLTSLVLGSSWCGGIP-----AIKPAISFSQRIYNGENAVLGSWPMQVSL 51
Db 1 MAFLLWLLSCWALLGTTFGCGVPDYKDDDDAAALAAPFDDDDKIVGVYALEAGQWPMQVSI 60
QY 52 QDSGPHFCGSLISQSWVYTAHCNVSPGRHF-----VVLGEYDRSSNAEPLOVLSVR 106
Db 61 -TYEGVHVCGLSVSEQWVLSAAHC--FPSEHHKEAYEVKLGALHQLDSYSEDAKAVSTLKD 117
QY 107 AIHPHSNNTMNDVTLLKLAPOYTRISPVCLASSNEALTEGLTCVTTGWGRLS-G 165
Db 118 IIPHSYLVQSGQDIALQLSRPITFSRYIRPICLPAAANSPNGLHCTVTGWHVAPS 177
QY 166 VGNVTPAHLOOVALPLVTYNOCROYWDSS-----ITDSMICAG--GAGASSCQGDGSG 215
Db 178 VSLITPKPLQOQLEVPCLISRETCLNYIDAKPEPHFVQDMVCAGYVEGGKDACQGDGSG 237
QY 216 GPLVCQKGNVTWVLIGI 231
Db 238 GPLSCPVEGLWYLTGI 253

```

Search completed: December 20, 2002, 15:16:39  
Job time : 37.6087 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:14:28 ; Search time 2.08696 Seconds  
(without alignments)  
1520.126 Million cell updates/sec

Title: US-09-856-319B-4\_COPY\_1\_33

Perfect score: 166

Sequence: 1 MLLSLTSLVLLGSSWCGVPAITPALSYNQR 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	150	90.4	264	2 I38136	chymotrypsin-like
2	76	45.8	263	2 A21195	chymotrypsin (EC 3
3	73	44.0	263	2 A31299	chymotrypsin (EC 3
4	67	40.4	263	1 KYRTB	chymotrypsin (EC 3
5	55.5	33.4	835	2 F90260	hypothetical prote
6	55	33.1	231	2 A49265	flit3/flk-2 ligand
7	55	33.1	242	2 AE0906	monofunctional bio
8	55	33.1	268	2 S68826	pancreatic elastas
9	55	33.1	268	2 S68825	pancreatic elastas
10	55	33.1	466	2 T05664	hypothetical prote
11	52	31.3	219	2 B71511	probable sugar nuc
12	52	31.3	241	2 A10431	monofunctional bio
13	52	31.3	292	2 T10106	chitinase (EC 3.2.
14	52	31.3	557	2 JE0346	high-affinity carn
15	52	31.3	579	2 T04378	peptide transport
16	51	30.7	826	2 T02268	potassium transpor
17	51	30.7	1186	2 T03180	tyrosine protein k
18	50.5	30.4	384	2 AH0241	ceramide glucosylt
19	50	30.1	247	2 S23504	chymase (EC 3.4.21
20	50	30.1	260	2 S26043	chymase (EC 3.4.21
21	50	30.1	363	1 CBUTB	ubiquinol-cytochro
22	50	30.1	742	2 S55098	probable membrane
23	49.5	29.8	312	1 Q0E007	US10 protein - hum
24	49.5	29.8	484	2 T10255	nitrite transport
25	49.5	29.8	739	1 A70204	endopeptidase Clp
26	49.5	29.8	918	2 F86709	cation-transportin
27	49	29.5	94	2 A83108	hypothetical prote
28	49	29.5	104	2 A22706	vitelline membrane
29	49	29.5	141	2 F72121	ATP synthase, chai

## ALIGNMENTS

### RESULT 1

I38136

chymotrypsin-like proteinase (EC 3.4.21.-) CTRL-1 - human

C:Species: Homo sapiens (man)

C>Date: 06-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change 22-Jun-1999

C:Accession: I38136

R:Larsen, F.; Solheim, J.; Kristensen, T.; Kolsto, A.B.; Prydz, H.

Hum. Mol. Genet. 2, 1589-1595, 1993

A:Title: A tight cluster of five unrelated human genes on chromosome 16q22.1.

A:Reference number: I38135; MUID:94093544; PMID:8268911

A:Accession: I38136

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-264 <RES>

A:Cross-references: EMBL:X71874; NID:g406226; PIDN:CAA50710.1; PID:g406228

C:Genetics:

A:Gene: GDB:CTRL

A:Cross-references: GDB:204061

A:Map position: 16q22.1-16q22.1

A:Introns: 18/1; 52/3; 79/2; 106/3; 167/1; 211/3

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; serine proteinase

F:34-257/Domain: trypsin homology <TRY>

F:75,121,214/Active site: His, Asp, Ser #status predicted

Query Match 90.4%; Score 150; DB 2; Length 264;

Best Local Similarity 87.9%; Pred. No. 9.2e-12;

Matches 29; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLLSLTSLVLLGSSWCGVPAITPALSYNQR 33

Db 1 MLLSLTSLVLLGSSWCGGIPAIKPALSFQR 33

### RESULT 2

A21195

chymotrypsin (EC 3.4.21.1) 2 precursor - dog

C:Species: Canis lupus familiaris (dog)

C>Date: 27-Jul-1990 #sequence\_revision 27-Jul-1990 #text\_change 22-Jun-1999

C:Accession: A21195

R:Pinsky, S.D.; LaForge, K.S.; Luc, V.; Scheele, G.

Proc. Natl. Acad. Sci. U.S.A. 80, 7486-7490, 1983

A:Title: Identification of cDNA clones encoding secretory isoenzyme forms: sequence d

A:Reference number: A21195; MUID:84170253; PMID:6584866

A:Accession: A21195

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-263 <PIN>

A:Cross-references: GB:K01173; NID:gl63945; PIDN:AAA30841.1; PID:gl63946

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; protein digestion; serine proteinase

F:34-256/Domain: trypsin homology <TRY>

1

Query Match 33.1%; Score 55; DB 2; Length 268;  
Best Local Similarity 40.7%;  
Pred. No. 12;  
Matches 11; Conservative 5; Mismatches 11; Indels 0;  
Gaps 0;

```

F;44-292/Product: chitinase #status experimental <MAT>
F;48-284/Domain: plant chitinase homology <PCH>

Query Match          31.3%; Score 52; DB 2; Length 292;
Best Local Similarity 55.0%; Pred. No. 30;
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLGSSWGCG 20
   : : | | | | | | :
Db 7 LLFSLVLSFVLGSAQNCG 26

RESULT 14
JEO346
high-affinity carnitine transporter, CTL - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: JEO346
R:Sekine, T.; Kusuhabara, H.; Utsunomiya-Tate, N.; Tsuda, M.; Sugiyama, Y.; Kanai, Y.;
Biochem. Biophys. Res. Commun. 251, 586-591, 1998
A:Title: Molecular cloning and characterization of high-affinity carnitine transporter
A:Reference number: JEO346; MUID:99011422; PMID:9792817
A:Accession: JEO346
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-557 <SEK>
A:Cross-references: DDBJ:AB017260; NID:93869208; PIDN:BAA34399.1; PID:93869209

Query Match          31.3%; Score 52; DB 2; Length 557;
Best Local Similarity 61.1%; Pred. No. 54;
Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLGSSWG 18
   | | | | | | : | | : | |
Db 258 MLLALTVPGLCALMG 275

RESULT 15
T04378
peptide transport protein - barley
C:Species: Hordeum vulgare (barley)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C:Accession: T04378
R:West, C.E.; Waterworth, W.M.; Stephens, S.M.; Smith, C.P.; Bray, C.M.
Plant J. 15, 221-229, 1998
A:Title: Cloning and functional characterisation of a peptide transporter expressed in
A:Reference number: Z15321; MUID:98388652; PMID:9721680
A:Accession: T04378
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-579 <WES>
A:Cross-references: EMBL:AF023472; NID:92655097; PIDN:AAC32034.1; PID:92655098
A:Experimental source: cv. Maris Otter
C:Genetics:
A:Gene: ptr1

Query Match          31.3%; Score 52; DB 2; Length 579;
Best Local Similarity 38.7%; Pred. No. 56;
Matches 12; Conservative 6; Mismatches 7; Indels 6; Gaps 1;

QY 5 SLTSLVL-----LGSSWGCGVPAITPALS 29
   : | | | | : | | | | | | : | | :
Db 196 ALVASSVLVTVQTHVGSWCGFIPAVVMAIA 226

Search completed: December 20, 2002, 15:19:49
Job time : 3.08696 secs

```

Search completed: December 20, 2002, 15:19:49  
Job time : 3.08696 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:12:58 ; Search time 9.04348 seconds  
(without alignments)  
1210.790 Million cell updates/sec

Title: US-09-856-319B-4

Perfect score: 1418

Sequence: 1 MLLSLTLVLVSSWGGG.....AMYTRVSKFSTWVQWYAYN 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1269	89.5	264	1	CTRL_HUMAN
2	832.5	58.7	263	1	CTRL_HUMAN
3	822.5	58.0	263	1	CTRL_HUMAN
4	798.5	56.3	263	1	CTRL_HUMAN
5	786.5	55.5	245	1	CTRL_HUMAN
6	781.5	55.1	263	1	CTRL_HUMAN
7	774.5	54.6	245	1	CTRL_HUMAN
8	748.5	52.8	245	1	CTRL_HUMAN
9	537	37.9	269	1	CTRL_HUMAN
10	520	36.7	269	1	CTRL_HUMAN
11	517	36.5	271	1	CTRL_HUMAN
12	515	36.3	269	1	CTRL_HUMAN
13	513.5	36.2	270	1	CTRL_HUMAN
14	503.5	35.5	253	1	CTRL_HUMAN
15	495.5	34.9	268	1	CTRL_HUMAN
16	492	34.7	271	1	CTRL_HUMAN
17	486.5	34.3	268	1	CTRL_HUMAN
18	485.5	34.2	270	1	CTRL_HUMAN
19	484.5	34.2	270	1	CTRL_HUMAN
20	484	34.1	810	1	CTRL_HUMAN
21	479	33.8	269	1	CTRL_HUMAN
22	477	33.6	343	1	CTRL_HUMAN
23	475	33.5	810	1	CTRL_HUMAN
24	474	33.4	790	1	CTRL_HUMAN
25	470	33.1	333	1	CTRL_HUMAN
26	468.5	33.0	342	1	CTRL_HUMAN
27	468	33.0	454	1	CTRL_HUMAN
28	465	32.8	270	1	CTRL_HUMAN
29	463.5	32.7	273	1	CTRL_HUMAN
30	462.5	32.6	342	1	CTRL_HUMAN
31	462.5	32.6	343	1	CTRL_HUMAN
32	460.5	32.5	275	1	CTRL_HUMAN
33	459	32.4	338	1	CTRL_HUMAN

34	454.5	32.1	246	1	TRY1_RAT	P00762 rattus norv
35	452.5	31.9	273	1	MCT7_RAT	P27435 rattus norv
36	451.5	31.8	490	1	TMS2_MOUSE	Q9198 mus musculu
37	447	31.5	231	1	TRYP_PIG	P00761 sus scrofa
38	447	31.5	266	1	EL1_BOVIN	Q28153 bos taurus
39	445.5	31.4	274	1	MCT6_RAT	P50343 rattus norv
40	444	31.3	266	1	EL1_RAT	P00773 rattus norv
41	443.5	31.3	275	1	TRY1_CANFA	P35944 canis faml
42	443	31.2	810	1	PLMN_ERIEU	Q29445 erinaceus e
43	442.5	31.2	346	1	TRY1_CANFA	P06871 canis faml
44	442	31.2	248	1	TRY1_CHICK	Q30627 gallus gall
45	441.5	31.1	812	1	PLMN_BOVIN	P06868 bos taurus

## ALIGNMENTS

### RESULT 1

ID	CTRL_HUMAN	STANDARD;	PRT;	264 AA.
AC	P40313;			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Chymotrypsin-like protease CTRL-1 precursor (EC 3.4.21.-).			
GN	CTRL OR CTRL1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=94093544; PubMed=8268911;			
RA	Larsen F., Solheim J., Kristensen T., Kolsto A.B., Prydz H.;			
RT	"A tight cluster of five unrelated human genes on chromosome			
RT	16q22.1.";			
RL	Hum. Mol. Genet. 2:1589-1595(1993).			
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X71874; CAA50710.1; -			
DR	EMBL; X71877; CAA50711.1; -			
DR	HSSP; P00763; LDPO.			
DR	MEROPS; S01_256; -			
DR	Genew; HGNC:2524; CTRL.			
DR	MIN; I18888; -			
DR	InterPro; IPR001314; Chymotrypsin.			
DR	InterPro; IPR001254; Ser-protease_Try.			
DR	Pfam; PF00089; trypsin; 1			
DR	PRINTS; PR00722; CHYMOTRYPSIN.			
DR	SMART; SM00020; TRY-SPC; 1.			
DR	PROSITE; PS50240; TRYPSIN_DOM; 1.			
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.			
DR	PROSITE; PS00135; TRYPSIN_SER; 1.			
KW	Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal.			
FT	SIGNAL	1	18	POTENTIAL.
FT	PROPEP	19	33	ACTIVATION PEPTIDE (POTENTIAL).
FT	CHAIN	34	264	CHYMOTRYPSIN-LIKE PROTEASE CTRL-1.
FT	ACT_SITE	75	75	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE	121	121	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE	214	214	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	CARBOHYD	114	114	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	19	141	BY SIMILARITY.
FT	DISULFID	60	76	BY SIMILARITY.
FT	DISULFID	155	220	BY SIMILARITY.
FT	DISULFID	187	201	BY SIMILARITY.

FT DISULFID 210 239 BY SIMILARITY.  
SQ SEQUENCE 264 AA; 28002 MW; 3FG29F02FA6DDFB4 CRC64;

Query Match 89.5%; Score 1269; DB 1; Length 264;  
Best Local Similarity 86.4%; Pred. No. 6.6e-102;  
Matches 228; Conservative 25; Mismatches 11; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLGSGGCGVPAITPALSYNQRIVNGENAVPGSWPQVSLQDNTGFHFC 60  
DB 1 MLLSLTSLVLGSGGCGVPAITPALSFSQRIVNGENAVPGSWPQVSLQDSSGFHFC 60

QY 61 GGSLSIPNVVTAHACQVTPGRHFVVLGGEYDRSSNAEPVQVLSIARAIHPNNMNTMN 120  
DB 61 GGSLSISQSWVYTAHACVSPGRHFVVLGGEYDRSSNAEPVQVLSIARAIHPNSWNTTMMN 120

QY 121 DLTLKLPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQOVVLP 180  
DB 121 DVTLKLPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQOVVLP 180

QY 181 LVTNOCROYWGARIITDAMICAGSGGCGSGGGLVCGKNTWVLIGIVSWGTFKNCN 240  
DB 181 LVTNOCROYWGARIITDAMICAGSGGCGSGGGLVCGKNTWVLIGIVSWGTFKNCN 240

QY 241 IQAPAMYTRVSKFSTWVNOVAYN 264  
DB 241 VRAPAVYTRVSKFSTWVNOVAYN 264

RESULT 2  
CTR2\_CANFA  
ID CTR2\_CANFA STANDARD; PRT; 263 AA.  
AC P04813;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Chymotrypsinogen 2 precursor (EC 3.4.21.1).  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=84170253; PubMed=6584866;  
RA Pinsky S.D., Laforce K.S., Luc V., Scheele G.;  
RT Identification of cDNA clones encoding secretory isoenzyme forms:  
RT sequence determination of canine pancreatic chymotrypsinogen 2 mRNA.  
RL Proc. Natl. Acad. Sci. U.S.A. 80:7486-7490(1983).  
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa, Phe-I-Xaa, Leu-I-Xaa.  
CC -1- SUBCELLULAR LOCATION: Extracellular.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; K01173; AAA30841.1; -  
CC DR PIR; A21195; A21195.  
CC DR HSSP; P00766; IACH.  
CC DR MEROPS; S01.152; -  
CC InterPro; IPR001314; Chymotrypsin.  
CC InterPro; IPR001254; Ser\_protease\_Try.  
CC Pfam; PF00089; trypsin; 1.  
CC PRINTS; PR00722; CHYMOTRYPSIN.  
CC SMART; SM00020; TRYP\_SPC; 1.  
CC PROSITE; PS0240; TRYP\_SIN\_DOM; 1.  
CC PROSITE; PS00134; TRYP\_SIN\_HTS; 1.  
CC PROSITE; PS00135; TRYP\_SIN\_SER; 1.

KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.  
FT SIGNAL 1 18  
FT CHAIN 19 263 CHYMOTRYPSINOGEN 2.  
FT CHAIN 19 31 CHYMOTRYPSIN 2, A CHAIN.  
FT CHAIN 34 164 CHYMOTRYPSIN 2, B CHAIN.  
FT CHAIN 157 263 CHYMOTRYPSIN 2, C CHAIN.  
FT ACT\_SITE 75 75 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 120 120 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 213 213 CHARGE RELAY SYSTEM.  
FT DISULFID 19 140 BY SIMILARITY.  
FT DISULFID 60 76 BY SIMILARITY.  
FT DISULFID 154 219 BY SIMILARITY.  
FT DISULFID 186 200 BY SIMILARITY.  
FT DISULFID 209 238 BY SIMILARITY.  
SQ SEQUENCE 263 AA; 27787 MW; 2A2F449D813B3961 CRC64;

Query Match 58.7%; Score 832.5; DB 1; Length 263;  
Best Local Similarity 57.2%; Pred. No. 1.9e-64;  
Matches 151; Conservative 38; Mismatches 74; Indels 1; Gaps 1;

QY 1 MLLSLTSLVLGSGGCGVPAITPALSYNQRIVNGENAVPGSWPQVSLQDNTGFHFC 60  
DB 1 MAFLLWLLSCFALLGTAFCGVPAPVLPVLSGLSRVNGEDAVPGSWPQVSLQDSTGFHFC 60

QY 61 GGSLSIPNVVTAHACQVTPGRHFVVLGGEYDRSSNAEPVQVLSIARAIHPNNMNTMN 120  
DB 61 GGSLSIEDWVTAHACGVRT-THQVVAGEFDQGSDAESIQVLKIAKVFKNPKENMTINN 119

QY 121 DLTLKLPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQOVVLP 180  
DB 121 DLTLKLPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQOVVLP 179

QY 181 LVTNOCROYWGARIITDAMICAGSGGCGSGGGLVCGKNTWVLIGIVSWGTFKNCN 240  
DB 180 LLSNAECKKFWGSKIITDLAVCAGASVSCMGDSGGGLVCGKDCAWTLVIGVSWGSGTCS 239

QY 241 IQAPAMYTRVSKFSTWVNOVAYN 264  
DB 240 TSTPGVYARVTKLIPWVQILQAN 263

RESULT 3  
CTR2\_HUMAN  
ID CTR2\_HUMAN STANDARD; PRT; 263 AA.  
AC P17538;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Chymotrypsinogen B precursor (EC 3.4.21.1).  
GN CTRB1 OR CTRB.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Pancreas;  
RX MEDLINE=89134264; PubMed=2917002;  
RA Tomita N., Izumoto Y., Horii A., Doi S., Yokouchi H., Ogawa M., Mori T., Matsubara K.;  
RA "Molecular cloning and nucleotide sequence of human pancreatic prechymotrypsinogen cDNA."  
RL Biochem. Biophys. Res. Commun. 158:569-575(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Pancreas;  
RL Strausberg R.;  
RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa, Phe-I-Xaa, Leu-I-Xaa.  
CC -1- SUBCELLULAR LOCATION: Extracellular.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
EMBL: M24400; AAA52128.1; --  
EMBL: BC005385; AAH05385.1; --  
PIR: A31299; A31299.  
HSP: P00766; ICHG.  
MEROPS: S01.152; --  
Genew: HGNC:2521; CTRB1.  
MIM: 118890; --  
InterPro: IPR001314; Chymotrypsin.  
InterPro: IPR001254; Ser\_protease\_Try.  
Pfam: PF00089; trypsin; 1.  
PRINTS: PR00722; CHYMOTRYPSIN.  
SMART: SM00020; Tryp\_Spc; 1.  
PROSITE: PS00240; TRYP\_SPC; 1.  
PROSITE: PS00134; TRYP\_SIN\_HIS; 1.  
PROSITE: PS00135; TRYP\_SIN\_SER; 1.  
Hydrolase: Serine protease; Digestion; Pancreas; Zymogen; Signal.  
FT SIGNAL 1 18  
FT CHAIN 19 263 CHYMOTRYPSIN B.  
FT CHAIN 19 31 CHYMOTRYPSIN B, A CHAIN.  
FT CHAIN 34 164 CHYMOTRYPSIN B, B CHAIN.  
FT CHAIN 167 263 CHYMOTRYPSIN B, C CHAIN.  
FT ACT\_SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT DISULFID 19 140 BY SIMILARITY.  
FT DISULFID 60 76 BY SIMILARITY.  
FT DISULFID 154 219 BY SIMILARITY.  
FT DISULFID 186 200 BY SIMILARITY.  
FT DISULFID 209 238 BY SIMILARITY.  
SQ SEQUENCE 263 AA; 27870 MW; 4C1C055A490B8701 CRC64;

Query Match 58.08; Score 822.5; DB 1; Length 263;  
Best Local Similarity 57.28; Pred. No. 1.4e-63;  
Matches 151; Conservative 38; Mismatches 74; Indels 1; Gaps 1;  
Qy 1 MLLSLTSLVLLGSSWCGVPAITPALSYNORIVNGENAVPGSWPQVSLQDNTGPHFC 60  
Db 1 MAFWLWSCWALLGTTFCGVPALHPVLSGLSRVINGEDAVPGSWPQVSLQDNTGPHFC 60  
Qy 61 GGSLISPNWVYTAARHCQVTPGRHFVVLGEYDRSSNAEPVQVLSIARATHPNWNTMNN 120  
Db 61 GGSLISEDVWVYTAARCGVRTS-DVVVAGEFDGSDDEENIQVLKIAKVPKPKFSILTNN 119  
Qy 121 DTLKLLKASPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQOVVLP 180  
Db 120 DITLLKATPARFSQTVSAVCLPSADDDFPAGTLCATTGWGKTRYNANKTPDKLQQAALP 179  
Qy 181 LVTYNOCROYAGARTDDAMICAGSGSSCGDGGPLVLCOKGNTWVLIGIVSGTKNCN 240  
Db 180 LLSNAECKSGWRRTDVMICAGASGVSSCMGDSGGPLVLCOKDGAWTLVGVSGSDTCS 239  
Qy 241 IQAPAMYTRVSKFSTWVQVMAYN 264  
Db 240 TSSPGVYARVTKLIPWQKILAN 263

RESULT 4  
ID CTRA\_GADMO STANDARD; PRT; 263 AA.  
AC P47796;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Chymotrypsin A precursor (EC 3.4.21.1).  
OS Gadus morhua (Atlantic cod).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorphina; Paracanthopterygii; Gadiformes; Gadidae; Gadus.  
RN NCBI\_TaxID=8049;  
RP [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Pyloric caeca;  
RX MEDLINE=94368860; PubMed=8086467;  
RA Gudmundsdottir A., Oskarsson S., Eakin A.E., Craik C.S.,  
RA Bjarnason J.B.;  
RT "Atlantic cod cDNA encoding a psychrophilic chymotrypsinogen.";  
RL Biochim. Biophys. Acta 1219:211-214(1994).  
[2]  
RP SEQUENCE OF 19-30 AND 34-49.  
RC TISSUE=Pyloric caeca;  
RX MEDLINE=92111252; PubMed=1764912;  
RA Ageirsson B., Bjarnason J.B.;  
RT "Structural and kinetic properties of chymotrypsin from Atlantic cod  
(Gadus morhua). Comparison with bovine chymotrypsin.";  
RL Comp. Biochem. Physiol. 99B:327-335(1991).  
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,  
CC Phe-I-Xaa, Leu-I-Xaa.  
CC -1- SUBCELLULAR LOCATION: Extracellular.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
EMBL: X78490; CAA55242.1; --  
HSP: P00766; ICHG.  
MEROPS: S01.152; --  
InterPro: IPR001314; Chymotrypsin.  
InterPro: IPR001254; Ser\_protease\_Try.  
Pfam: PF00089; trypsin; 1.  
PRINTS: PR00722; CHYMOTRYPSIN.  
SMART: SM00020; Tryp\_Spc; 1.  
PROSITE: PS00240; TRYP\_SPC; 1.  
PROSITE: PS00134; TRYP\_SIN\_HIS; 1.  
PROSITE: PS00135; TRYP\_SIN\_SER; 1.  
Hydrolase: Serine protease; Digestion; Pancreas; Zymogen; Signal.  
FT SIGNAL 1 18  
FT CHAIN 19 263 CHYMOTRYPSIN A.  
FT ACT\_SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT DISULFID 19 140 BY SIMILARITY.  
FT DISULFID 60 76 BY SIMILARITY.  
FT DISULFID 154 219 BY SIMILARITY.  
FT DISULFID 186 200 BY SIMILARITY.  
FT DISULFID 209 238 BY SIMILARITY.  
FT CONFLICT 21 21 R -> S (IN REF. 2).  
FT CONFLICT 25 25 S -> Q (IN REF. 2).  
FT CONFLICT 29 29 T -> S (IN REF. 2).  
FT CONFLICT 44 44 S -> T (IN REF. 2).  
FT CONFLICT 46 46 S -> Y (IN REF. 2).  
SQ SEQUENCE 263 AA; 28294 MW; 47AAC699A0A64FBB CRC64;

Query Match 56.38; Score 798.5; DB 1; Length 263;  
Best Local Similarity 56.98; Pred. No. 1.6e-61;  
Matches 144; Conservative 38; Mismatches 70; Indels 1; Gaps 1;  
Qy 12 LLGSSWCGVPAITPALSYNORIVNGENAVPGSWPQVSLQDNTGPHFCGSLISPNWV 71  
Db 12 LFRRTYGGCRPAISVITGYSRVINGEEAVPHSWQVSLQDQTFHFCGSLINWV 71  
Qy 72 TAAHCQVTPGRHFVVLGEYDRSSNAEPVQVLSIARATHPNWNTMNNNDTLKLLASPA 131  
Db 72 TAAHCNV-KNHYRVVLGEHRRSSNSEGQVQVMTVGQVFXPRNGFTINNDILLVKLATPA 130

QY 132 RYTAQVSPVCLASTNEALPSGLTCVTGWRISGVGNVTPARLQVVLPLVTNOCROYW 191  
 Db 131 TLNMRVSPVCLASTDDVFEKKKCVTSGLTFYNNADTFALLOALPLLTNEQCKFW 190  
 QY 192 GARITDAMICAGGASSCGGSGPLVCOGNWTLGIVSGTRKNCNIQAPAMVTRYS 251  
 Db 191 GNKISDLMICAGAGASSCGGSGPLVCOGNWTLGIVSGTRKNCNIQAPAMVTRYS 250  
 QY 252 KFTSWINQVWAYN 264  
 Db 251 ELRAWVDQITAAAN 263

RESULT 5  
 CTRB\_BOVIN STANDARD; PRT; 245 AA.  
 AC P00767;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Chymotrypsinogen B (EC 3.4.21.1).  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE, DISULFIDE BONDS, AND ACTIVE SITE.  
 RX MEDLINE=68238908; PubMed=5649671;  
 RA "Strillie L.B., Furka A., Nagabushan N., Stevenson K.J., Parkes C.O.; D. *Qued*  
 RT "Structure of chymotrypsinogen B compared with chymotrypsinogen A and  
 RT trypsinogen."  
 RL Nature 218:343-346(1968).  
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,  
 CC -1- SUBCELLULAR LOCATION: Extracellular.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -1- DATABASE: NAME=worthington enzyme manual;  
 CC WWW="http://www.worthington-biochem.com/manual/C/CHY.html".  
 DR PIR: A00953; KYBOB.  
 DR HSSP: P00766; IACB.  
 DR MEROPS: S01.152; -.  
 DR InterPro: IPR001314; Chymotrypsin.  
 DR InterPro: IPR001254; Ser\_protease\_Try.  
 DR Pfam: PF00089; trypsin; 1.  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR SMART: SM00020; TRYPSIN\_DOM; 1.  
 DR PROSITE: PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE: PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
 KW Hydroxylase; Serine protease; Digestion; Pancreas; Zymogen.  
 FT CHAIN 1 13 CHYMOTRYPSIN B, A CHAIN.  
 FT CHAIN 15 146 CHYMOTRYPSIN B, B CHAIN.  
 FT CHAIN 149 245 CHYMOTRYPSIN B, C CHAIN.  
 FT ACT\_SITE 57 57 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 102 102 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 195 195 CHARGE RELAY SYSTEM.  
 FT DISULFID 1 122  
 FT DISULFID 42 58  
 FT DISULFID 136 201  
 FT DISULFID 168 182  
 FT DISULFID 191 220  
 SQ SEQUENCE 245 AA; 25755 MW; 678016446FF5FEB5 CRC64;

Query Match 55.5%; Score 786.5; DB 1; Length 245;  
 Best Local Similarity 57.7%; Pred. No. 1.6e-60;  
 Matches 142; Conservative 32; Mismatches 71; Indels 1; Gaps 1;  
 QY 19 CGVPALTPALSNORLVNGENAVPGSWPQVSLQDNTGFHFCGGLSISPNWVYTAACOV 78  
 Db 1 CGVPALTPALSNORLVNGENAVPGSWPQVSLQDNTGFHFCGGLSISPNWVYTAACOV 78

QY 79 TPGREHVVILGEYDRSSNABPQVLSIARALTHPHNNANTMNNDLTLKLASPARYTAQVS 138  
 Db 61 TTS-DVVVAGEEDQLEDTQVLKIGKVKFKPFSILTVRNDITLKLATPAQSETVS 119  
 QY 139 PVLASTNEALPSGLTCVTGWRISGVGNVTPARLQVVLPLVTNOCROYWGARITDA 198  
 Db 120 AVCLPSADEDFPAGMLCATGKTKYNALKTDPKLLQATLPVSNVTDCKRYGSRVTDV 179  
 QY 199 MICAGSGASSCGGSGPLVCOGNWTLGIVSGTRKNCNIQAPAMVTRYSKFTWIN 258  
 Db 180 MICAGSGVSSCMGDSGGLVCOGNWTLGIVSGTRKNCNIQAPAMVTRYSKFTWIN 239  
 QY 259 QVMAYN 264  
 Db 240 ETALAN 245  
 RESULT 6  
 CTRB\_RAT STANDARD; PRT; 263 AA.  
 AC P07338;  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 01-APR-1988 (Rel. 07, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Chymotrypsinogen B precursor (EC 3.4.21.1).  
 GN CTRB1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85054881; PubMed=6209274;  
 RA Bell G.I., Quinto C., Quiroga M., Valenzuela P., Craik C.S.,  
 RA Rutter W.J.;  
 RT "Isolation and sequence of a rat chymotrypsin B gene."  
 RL J. Biol. Chem. 259:14265-14270(1984).  
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,  
 CC -1- SUBCELLULAR LOCATION: Extracellular.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL: K02298; AAA98732.1; -.  
 DR PIR: A22658; KYRTR.  
 DR HSSP: P00766; 1CHG.  
 DR MEROPS: S01.152; -.  
 DR InterPro: IPR001314; Chymotrypsin.  
 DR InterPro: IPR001254; Ser\_protease\_Try.  
 DR Pfam: PF00089; trypsin; 1.  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR SMART: SM00020; TRYPSIN\_DOM; 1.  
 DR PROSITE: PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE: PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
 KW Hydroxylase; Serine protease; Digestion; Pancreas; Zymogen; Signal.  
 FT SIGNAL 1 18  
 FT CHAIN 19 263 CHYMOTRYPSIN B.  
 FT CHAIN 19 31 CHYMOTRYPSIN B, A CHAIN.  
 FT CHAIN 34 164 CHYMOTRYPSIN B, B CHAIN.  
 FT CHAIN 167 263 CHYMOTRYPSIN B, C CHAIN.  
 FT ACT\_SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT DISULFID 19 140 BY SIMILARITY.  
 FT DISULFID 60 76 BY SIMILARITY.



```
FT DISULFID 154 219 BY SIMILARITY.
FT DISULFID 186 200 BY SIMILARITY.
FT DISULFID 209 238 BY SIMILARITY.
SQ SEQUENCE 263 AA; 27849 MW; ACAFDACF8C4DA6D CRC64;

Query Match 55.1%; Score 781.5; DB 1; Length 263;
Best Local Similarity 53.8%; Pred. No. 4.6e-60;
Matches 142; Conservative 40; Mismatches 81; Indels 1; Gaps 1;

QY 1 MLLSLTSLVLLSGGCGGPAITPALSYNQRIYNGENAVPGSWPQVSVLQDNTGFHFC 60
DB 1 MAFLWLVSCFALVGATGCGVPTIQVLTLGLSRIVNGDAIPGSPWQVSVLQDNTGFHFC 60
QY 61 GGSLSIPNWWYTAACQVTPGRHFEVVLGEXDRSSNAEPQVLSIARATHTPNWNAANTWNN 120
DB 61 GGSLSIEDWVYTAACGKVTSDVWVAGEFFDGSDEENIQVLKIAQVFNKFNKFNFTVRN 119
QY 121 DITLLKASPARYTAQVSPVCLASTNEALPSGLTCTVTGWRISGVNVTFARLQOVVLP 180
DB 121 DITLLKATPAQFSETVSACVLPNVDPPPGTVCAITGKTKYNALKTPEKIQQAALP 179
QY 181 LVTYNQCRQYWCARITDAMICAGSGASSCGQDGGPLVCQKQNTWVLIGIVSWGTKNKN 240
DB 181 IVSEADCKKSGSKITDYMTCCAGASGVSSCMGDSGGPLVCQKQDGYWTLGIVSWGSGVCS 239
QY 241 IQAPAMYTRVSKFSTWVINOVMAYN 264
DB 240 TSTPAVYSRVTAIMPWVOOILEAN 263

RESULT 7
CTRA_BOVIN STANDARD; PRT; 245 AA.
ID AC P00766;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymotrypsinogen A (EC 3.4.21.1).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE, DISULFIDE BONDS, AND ACTIVE SITE.
RX MEDLINE=67181721; PubMed=5971783;
RA Brown J.R., Hartley B.S.;
RT "Location of disulphide bridges by diagonal paper electrophoresis.
RT The disulphide bridges of bovine chymotrypsinogen A.";
RL Biochem. J. 101:214-228(1966).
RN [2]
RP REVISION TO 102.
RX MEDLINE=69106266; PubMed=5764436;
RA Blow D.M., Birktoft J.J., Hartley B.S.;
RT "Role of a buried acid group in the mechanism of action of
RT chymotrypsin.";
RL Nature 221:337-340(1969).
RN [3]
RP PRELIMINARY SEQUENCE.
RA Hartley B.S.;
RT "Amino-acid sequence of bovine chymotrypsinogen-A.";
RL Nature 201:1284-1287(1964).
RN [4]
RP SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=67183948; PubMed=5972866;
RA Meloun B., Klueh I., Kostka V., Moravec L., Prusik Z., Vanacek J.,
RA Kell B., Sorm F.;
RT "Covalent structure of bovine chymotrypsinogen A.";
RL Biochim. Biophys. Acta 130:543-546(1966).
RN [5]
RP ACTIVE SITE.
RX MEDLINE=67181723; PubMed=5971785;
RA Smillie L.B., Hartley B.S.;
```

```
RT "Histidine sequences in the active centres of some 'serine'
RT proteinases.";
RL Biochem. J. 101:232-241(1966).
RN [6]
RP X-RAY CRYSTALLOGRAPHY.
RX MEDLINE=72035052; PubMed=4399050;
RA Birktoft J.J., Blow D.M., Henderson R., Steitz T.A.;
RT "I. Serine proteinases. The structure of alpha-chymotrypsin.";
RL Philos. Trans. R. Soc. Lond., B, Biol. Sci. 257:67-76(1970).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF CHYMOTRYPSINOGEN.
RX MEDLINE=70177557; PubMed=5442169;
RA Freer S.T., Kraut J., Robertus J.D., Wright H.T., Xuong N.H.;
RT "Chymotrypsinogen: 2.5-A crystal structure, comparison with alpha-
RT chymotrypsin, and implications for zymogen activation.";
RL Biochemistry 9:1997-2009(1970).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF GAMMA-CHYMOTRYPSIN.
RX MEDLINE=82078042; PubMed=6914398;
RA Cohen G.H., Silverton E.W., Davies D.R.;
RT "Refined crystal structure of gamma-chymotrypsin at 1.9-A resolution.
RT Comparison with other pancreatic serine proteases.";
RL J. Mol. Biol. 184:449-479(1981).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (1.68 ANGSTROMS) OF ALPHA-CHYMOTRYPSIN.
RX MEDLINE=86011575; PubMed=4046030;
RA Tsukada H., Blow D.M.;
RT "Structure of alpha-chymotrypsin refined at 1.68-A resolution.";
RL J. Mol. Biol. 184:703-711(1985).
CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,
CC Phe-I-Xaa, Leu-I-Xaa.
CC -|- SUBCELLULAR LOCATION: Extracellular.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -|- DATABASE: NAME=Worthington enzyme manual;
CC WWW="http://www.worthington-biochem.com/manual/C/CHY.html".
PIR; A00952; KYBOA.
DR PDB; 2CGA; 15-APR-90.
DR PDB; 2CHA; 31-MAY-84.
DR PDB; 4CHA; 29-OCT-85.
DR PDB; 5CHA; 16-OCT-87.
DR PDB; 6CHA; 16-OCT-87.
DR PDB; 1CHG; 27-JAN-84.
DR PDB; 1CHO; 16-JUL-88.
DR PDB; 2GCH; 31-MAY-84.
DR PDB; 3GCH; 15-OCT-92.
DR PDB; 4GCH; 15-OCT-90.
DR PDB; 5GCH; 15-OCT-90.
DR PDB; 6GCH; 15-OCT-90.
DR PDB; 7GCH; 15-OCT-90.
DR PDB; 8GCH; 15-JUL-93.
DR PDB; 1GCT; 15-OCT-91.
DR PDB; 2GCT; 15-OCT-91.
DR PDB; 3GCT; 15-OCT-91.
DR PDB; 1ACB; 31-OCT-93.
DR PDB; 1GMD; 31-OCT-93.
DR PDB; 1GMD; 31-OCT-93.
DR PDB; 1CGI; 30-APR-94.
DR PDB; 1CGJ; 22-JUN-94.
DR PDB; 1GHB; 22-JUN-94.
DR PDB; 1GMB; 30-SEP-94.
DR PDB; 2GMT; 01-NOV-94.
DR PDB; 1MTN; 17-AUG-96.
DR PDB; 1AFQ; 17-SEP-97.
DR PDB; 1CAO; 23-JUL-97.
DR PDB; 1GBW; 23-JUL-97.
DR PDB; 1VGC; 12-NOV-97.
DR PDB; 2VGC; 12-NOV-97.
DR PDB; 3VGC; 12-NOV-97.
DR PDB; 4VGC; 12-NOV-97.
DR PDB; 1HJA; 14-JAN-98.
```



```
FT CONFLICT 9 11 QVT -> VIS (IN REF. 2).
FT CONFLICT 26 26 S -> T (IN REF. 2).
FT CONFLICT 28 29 PW -> Y (IN REF. 2).
SQ SEQUENCE 245 AA; 26260 MW; 74FE0D425517AB02 CRC64;

Query Match 52.8%; Score 748.5; DB 1; Length 245;
Best Local Similarity 54.6%; Pred. No. 2.9e-57;
Matches 136; Conservative 43; Mismatches 63; Indels 7; Gaps 5;

QY 19 CGVPAITPALSYNORINGENAVPGSPWQVSLQDNTGFHFCGSLSPNWKVVTAAHCQV 78
Db 1 CGSPAIOQVTVGYARRINGEAEVPHSPWQVSLQDNTGFHFCGSLSPNWKVVTAAHCN 60
QY 79 TPGHFVVLGYDRSSNAEPQVLSIARAIHPNNMNDLTLKLASPARYTQVS 138
Db 61 RT-VHRVIGEHKASD-ENIQILKPSVTFHPKWDERTINDISLKLASPAVLCTNVS 118
QY 139 PVLASTNEALPSGLTCTVTGWS--RISGVGNVTPARLQVPLVTNQCQVWGARG-I 195
Db 119 PVLGESSDVFPAGMKCVTSWGLTRYNAPG--TPNKLQQAALPLMSNEECSTWGNMI 176
QY 196 TDMICAGGSCASSCGDSGGLPCVCKGNTWVLGIYSWGTKNICIQAPAMYTRVSKEST 255
Db 177 SDVMICAGAAGATSCMGDSGGLPCVCKQDNVMTLGIYSWSSRCVSVTPAVYARVTLRG 236
QY 256 WINQVMAYN 264
Db 237 WVDQILAN 245

RESULT 9
EL2_PIG STANDARD; PRT; 269 AA.
ID EL2_PIG
AC P08419;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Elastase 2 precursor (EC 3.4.21.71).
GN ELA2.
OS Sus scrofa (Pig).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=67217962; PubMed=3646943;
RA Kawashima I., Tani T., Shimoda K., Takiguchi Y.;
RT "Characterization of pancreatic elastase II cDNAs: two elastase II
RL mRNAs are expressed in human pancreas."
RL DNA 6:163-172(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88198076; PubMed=2834346;
RA Shirasu Y., Yoshida H., Matsuki S., Takemura K., Ikeda N.,
RA Shimada Y., Ozawa T., Mikayama T., Iijima H., Ishida A., Sato Y.,
RA Tamai Y., Tanaka J., Ikenaga H.;
RT "Molecular cloning and expression in Escherichia coli of a cDNA
RL encoding human pancreatic elastase 2."
RL J. Biochem. 102:1555-1563(1987).
CC -I- FUNCTION: ACTS UPON ELASTIN.
CC -I- CATALYTIC ACTIVITY: Preferential cleavage: Leu-I-Xaa, Met-I-Xaa
CC and Phe-I-Xaa. Hydrolyzes elastin.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- TISSUE SPECIFICITY: PANCREAS.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC -----
EMBL: M16651; AAA31027.1; -.
DR EMBL: D00237; BAA00166.1; -.
DR PIR: A26823; A26823.
DR HSSP: P00763; IDPO.
DR MEROPS: S01.155; -.
DR InterPro: IPR001314; Chymotrypsin.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; Tryp_spc; 1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Zymogen; Signal.
FT SIGNAL 1 16
FT PROPEP 17 28 ACTIVATION PEPTIDE.
FT CHAIN 29 269 ELASTASE 2.
FT ACT_SITE 73 73 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 121 121 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 216 216 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 58 74 BY SIMILARITY.
FT DISULFID 155 222 BY SIMILARITY.
FT DISULFID 186 202 BY SIMILARITY.
FT DISULFID 212 243 BY SIMILARITY.
FT CONFLICT 10 10 L -> S (IN REF. 2).
FT CONFLICT 118 118 N -> K (IN REF. 2).
FT CONFLICT 132 132 S -> Y (IN REF. 2).
FT CONFLICT 172 172 I -> V (IN REF. 2).
FT CONFLICT 202 202 C -> V (IN REF. 2).
SQ SEQUENCE 269 AA; 28699 MW; BAC6FE69AF4DDE56 CRC64;

Query Match 37.9%; Score 537; DB 1; Length 269;
Best Local Similarity 43.1%; Pred. No. 4.5e-39;
Matches 118; Conservative 48; Mismatches 92; Indels 16; Gaps 11;

QY 2 LLLSLTSLVLLGSSWGGVPAITPALSYNORINGENAVPGSPWQVSLQ-DNTG--FH 58
Db 1 MIRALLSTLVAG-ALSCGLPANLPQL---PRVVGEDARPNSWQVSLQYDSSGQWRH 56
QY 59 FCGSLSPNWKVVTAAHCQVTPGRHFVVLGEYDRSSNAEPVQVLSIARAIHPNNMNTM 118
Db 57 TCGGTLVDQSWLTAACHICISSRTYRVVLRHSLSTNEPGLSKVSLVVRHQDWSNQL 116
QY 119 N--NDLTLLKLASPARYTAQVSPVCLASTNEALPSGLTCTVTGMRISGVGNVTPARLQ 176
Db 117 SNGNDIALKLASPVSLTDRQLGLCPAAGTILPNYVYVTGMRLOTNG-ASPDILQ 175
QY 177 VVLPVTVNQCQ--YNGARITDMICAGSG-ASSCGDSGGPLVCQKGN-TWVLIGIV 232
Db 176 GOLLVVDYATCSKPGWMSGSTVKTNNICAGDGIISCCNGDSGGPLNCQCGQMVHGIV 235
QY 233 SWGTR-NCN-IQAPAMYTRVSKFSTWVNOVMAYN 264
Db 236 SFGSSLCGNCYHKPSVFETRVSNYIDWINSVIANN 269

RESULT 10
EL2_BOVIN STANDARD; PRT; 269 AA.
ID EL2_BOVIN
AC Q29461;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Elastase 2 precursor (EC 3.4.21.71).
GN ELA2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
```

```

RC RX MEDLINE=98079203; PubMed=9418008;
RA Gestin M., le Huerou-Luron I., Wicker-Planquart C., le Drean G.,
RA Chai J.C., Puigserver A., Guilloteau P.
RT "Bovine pancreatic preproelastases I and II: comparison of nucleotide
RT and amino acid sequences and tissue specific expression.";
RL Comp. Biochem. Physiol. 118B:181-187(1997).
CC -!- FUNCTION: ACTS UPON ELASTIN.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Leu|-Xaa, Met|-Xaa
CC and Phe|-Xaa. Hydrolyzes elastin.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: PANCREAS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X97635; CAA66231.1; -.
DR HSSP; P00766; LCHG.
DR MEROPS; S01.155; -.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRY-SPC; 1.
DR PROSITE; PS0240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hydrolase; Serine protease; Zymogen; Signal.
FT SIGNAL 1 16
FT PROPEP 17 28 ACTIVATION PEPTIDE (BY SIMILARITY).
FT CHAIN 29 269 ELASTASE 2.
FT ACT_SITE 73 73 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 121 121 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 216 216 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 58 74 BY SIMILARITY.
FT DISULFID 155 222 BY SIMILARITY.
FT DISULFID 186 202 BY SIMILARITY.
FT DISULFID 212 243 BY SIMILARITY.
SQ SEQUENCE 269 AA; 28856 MW; 8343B97062CF267C CRC64;

Query Match 36.7%; Score 520; DB 1; Length 269;
Best Local Similarity 42.0%; Pred. No. 1.3e-37;
Matches 115; Conservative 45; Mismatches 98; Indels 16; Gaps 10;

QY 2 LLLSLTSLVLLGSSGCGVPAITPALSYNQRIVNCENAVPGSWPQVSLQDNTG---PH 58
Db 1 MIRALLSLTVAG-AUSCGVPTYPPLS---RVVGEDARPNWPQVSLQYSSSCQWRH 56

QY 59 FCGSLISPNWVVAACHQVTPGRHFEVLGEXDRSSNAEPVQVLTARATHFPNNNANTM 118
Db 57 TCGGSLIEQNWLTAAHCISSTRTYRVVVGROSLFVSSESLTIAVSKSVIHEKWNQNL 116

QY 119 --NNDTLKLKASPARTAQVSPVCLASTNEALPSGLTCVTGWRISGVGNVTPARLQ 176
Db 117 AQGNDAIALLKLASSVPLTDKIQGLCLPAAAGTILPNNVYCVYTWGRLQSG-ALPDILQ 175

QY 177 VVLPVTVNQCR--QYWGARTIDAMICAGSG--ASSCGDSDGGLVCQKGN-TWVLIGIV 232
Db 176 GLLVVDYATCSNPWSWGTVTKNIMCAGDGVTSNCNDSGGPLNCQANKQWQVHGIV 235

QY 233 SWGTK-NCN-IQAPAMYTRVSKFSTWVINOVMAYN 264
Db 236 SFGSSGLGCNYRKPSVTRVSNVNDWISSVIENN 269

RESULT 11
EL2_RAT
ID AC P00774;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Elastase 2 precursor (EC 3.4.21.71).
GN ELA2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82182967; PubMed=6918221;
RA McDonald R.J., Swift G.H., Quinto C., Swain W., Pictet R.L.,
RA Nikovits W., Rutter W.J.;
RT "Primary structure of two distinct rat pancreatic preproelastases
RT determined by sequence analysis of the complete cloned messenger
RT ribonucleic acid sequences.";
RL Biochemistry 21:1453-1463(1982).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85054882; PubMed=6094548;
RA Swift G.H., Craik C.S., Stary S.J., Quinto C., Lahaie R.G.,
RA Rutter W.J., Macdonald R.J.;
RT "Structure of the two related elastase genes expressed in the rat
RT pancreas.";
RL J. Biol. Chem. 259:14271-14278(1984).
CC -!- FUNCTION: ACTS UPON ELASTIN.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Leu|-Xaa, Met|-Xaa
CC and Phe|-Xaa. Hydrolyzes elastin.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: PANCREAS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V01233; CAA24543.1; -.
DR EMBL; L00124; AAA98780.1; -.
DR EMBL; L00118; AAA98780.1; JOINED.
DR EMBL; L00119; AAA98780.1; JOINED.
DR EMBL; L00120; AAA98780.1; JOINED.
DR EMBL; L00121; AAA98780.1; JOINED.
DR EMBL; L00122; AAA98780.1; JOINED.
DR EMBL; L00123; AAA98780.1; JOINED.
DR PIR; A00961; ELRT2.
DR HSSP; P00772; IELG.
DR MEROPS; S01.155; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRY-SPC; 1.
DR PROSITE; PS0240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hydrolase; Serine protease; Zymogen; Signal.
FT SIGNAL 1 16
FT PROPEP 17 30 ACTIVATION PEPTIDE.
FT CHAIN 31 271 ELASTASE 2.
FT ACT_SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 123 123 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 218 218 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 60 76 BY SIMILARITY.
FT DISULFID 157 224 BY SIMILARITY.
FT DISULFID 188 204 BY SIMILARITY.
FT DISULFID 214 245 BY SIMILARITY.

```

-|- CATALYTIC ACTIVITY: Preferential cleavage: Leu-|-Xaa, Met-|-Xaa  
and Phe-|-Xaa. Hydrolyzes elastin.  
-|- SUBCELLULAR LOCATION: Secreted.  
-|- TISSUE SPECIFICITY: PANCREAS.  
-|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute.. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to licensee@isb-sib.ch).

EMBL; M16631; AAA52374.1; -.  
DR EMBL; M16652; AAA52380.1; -.  
DR EMBL; D00236; BAA00165.1; -.  
DR EMBL; AL512883; CAC42421.1; -.  
DR EMBL; BC007031; AAH07031.1; -.  
DR PIR; A27432; A27432.  
DR PIR; B26823; B26823.  
DR HSSP; P00772; IELG.  
DR MEROPS; S01\_155; -.  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR001254; Ser\_protease\_Try.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; TRY\_PSPC; 1.  
DR PROSITE; PS00240; TRYP\_SIN\_DOM; 1.  
DR PROSITE; PS00134; TRYP\_SIN\_HIS; 1.  
DR PROSITE; PS00135; TRYP\_SIN\_SER; 1.  
KW Hydrolase; Serine protease; zymogen; Signal.  
FT SIGNAL 1 16  
FT FT PROPEP 17 28 ACTIVATION PEPTIDE.  
FT CHAIN 29 269  
FT DISULFID 58 74  
FT DISULFID 155 222 BY SIMILARITY.  
FT DISULFID 186 202 BY SIMILARITY.  
FT DISULFID 212 243 BY SIMILARITY.  
FT ACT\_SITE 73 73 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 121 121 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 216 216 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT CONFLICT 202 202 C -> V (IN REF. 3).  
SQ SEQUENCE 269 AA; 2888 MW; A2E05143EFF4987C CRC64;

Query Match :36.3%; Score 515; DB 1: Length 269;  
Best Local Similarity 41.6%; Pred.No 3.5e-37;  
Matches 114; Conservative 47; Mismatches 97; Indels 16; Gaps 10;

QY 2 LLLSLTLSTLVLLSGSGCGVPAITPALSYNQRIVNGENAVPGSNPMQVSLODNTG---FH 58  
: :: |::| :: || |::||| ::|::|::|::|::|::|:  
Db 1 MIRTLTLSTLVAG-ALSCGDTPPP--YYTRVVGGEARPNKSWPMQVSLQYSNGWKWH 56

QY 59 FCGSSLIPNNVTAAHCQVTPGRHFVLGEYDRSSNAEPQVLSIARAITHPNWNANTM 118  
|||||: |:|:||||| ::|::|::|::|::|::|:  
Db 57 TCQGSLIANSVWLTAACHSISSRTYRGVGLRHNLHYAESGLVASVSKIVVHKDWNSNQI 116

QY 119 N-NDLTKLKASPARYAQSVPVCLASTNEALPSGLTCVTGTGWGISGVGNVTPARQQ 176  
: |:|:|||||:|::|::|::|::|::|::|::|::|::|::|:  
Db 117 SKGNDIALKLAMPVSUTDKRIACLDPACTILTLPNNNYCYVVGMRLOTNGAV-PDVLIQ 175

QY 177 VWPLTVTNQCRO--YGGARITDMACIGGSAGA-SSCQDGSGGPLVCQKN-TTWLTIGV 232  
|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Db 176 GRLLVDVDTYCSSAWWGWSVKTSMICAGDGYISSCNCGDSGGLNCQASDGRWQHGV 235

QY 233 SWGTK-NCN-IQAPAMYTRYSKFSWTWINQNVAYN 264  
|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Db 236 SPFGSRLCGNCYHHKPVSFTRYSNYIDINSVIANN 269

RESULT 13  
EL3B\_HUMAN STANDARD; PROT; 270 AA.  
ID EL3B\_HUMAN

P08861; P11423;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Elastase IIIB precursor (EC 3.4.21.70) (Protease E).  
 GN ELA3B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Pancreas;  
 RX MEDLINE=88087253; PubMed=2826474;  
 RA Tani T., Ohsumi J., Mita K., Takiguchi Y.;  
 RT "Identification of a novel class of elastase isozyme, human  
 RT pancreatic elastase III, by cDNA and genomic gene cloning.";  
 RL J. Biol. Chem. 263:1231-1239(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Pancreas;  
 RA Strausberg R.;  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 4-270 FROM N.A.  
 RC TISSUE=Pancreas;  
 RX MEDLINE=88000845; PubMed=3477287;  
 RA Shen W., Fletcher T.S., Largman C.;  
 RT "Primary structure of human pancreatic protease E determined by  
 RT sequence analysis of the cloned mRNA.";  
 RL Biochemistry 26:3447-3452(1987).  
 RN [4]  
 RP SEQUENCE OF 31-50.  
 RC TISSUE=Pancreas;  
 RX MEDLINE=89325560; PubMed=2753124;  
 RA Moulard M., Kerfelec B., Mallet B., Chapus C.;  
 RT "Identification of a procarboxypeptidase A-truncated protease E  
 RT binary complex in human pancreatic juice.";  
 RL FEBS Lett. 250:166-170(1989).  
 RN [5]  
 RP SEQUENCE OF 94-164, AND CARBOHYDRATE-LINKAGE SITES ASN-114.  
 RC TISSUE=Pancreas;  
 RX MEDLINE=89289996; PubMed=2737288;  
 RA Wendorf P., Geyer R., Szlegoleit A., Linder D.;  
 RT "Localization and characterization of the glycosylation site of human  
 RT pancreatic elastase 1.";  
 RL FEBS Lett. 249:275-278(1989).  
 CC -!- FUNCTION: EFFICIENT PROTEASE WITH ALANINE SPECIFICITY BUT ONLY  
 CC LITTLE ELASTOLYTIC ACTIVITY.  
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Ala-I-Xaa. Does not  
 CC hydrolyse elastin.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.  
 CC -!- CAUTION: Was originally (Ref.5) thought to be elastase 1.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M16630; AAA36482.1;  
 CC EMBL; BC005216; AAH05216.1;  
 CC EMBL; M18692; AAA58434.1;  
 CC PIR; B29934; B29934.  
 CC PIR; A27206; A27206.  
 CC PIR; S04999; S04999.  
 CC PIR; S04490; S04490.  
 CC HSP; P05805; 1FON.  
 CC MEROPS; S01.205;  
 CC GlycosultebB; P08861;  
 CC SWISS-2DPAGE; P08861; HUMAN.

DR Genew; HGNC:15945; ELA3B.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PRO0722; CHYMOTRYPsin.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS00134; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Hydrolase; Serine protease; Zymogen; Signal; Glycoprotein.  
 FT SIGNAL 1 15  
 FT OR 16 (POTENTIAL).  
 FT PROPEP 16 28  
 FT CHAIN 29 270  
 FT ELASTASE IIIB.  
 FT ACT\_SITE 73 73  
 FT CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 123 123  
 FT CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 217 217  
 FT CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT DISULFID 58 74  
 FT PROBABLE.  
 FT DISULFID 117 120  
 FT BY SIMILARITY.  
 FT DISULFID 157 223  
 FT BY SIMILARITY.  
 FT DISULFID 188 204  
 FT BY SIMILARITY.  
 FT DISULFID 213 244  
 FT BY SIMILARITY.  
 FT CARBOHYD 114 114  
 FT N-LINKED (GLCNAC. . .).  
 FT /FTIQ-CAR\_000212.  
 FT CONFLICT 4 4  
 FT R -> G (IN REF. 3).  
 FT CONFLICT 64 64  
 FT A -> G (IN REF. 3).  
 FT CONFLICT 79 79  
 FT W -> R (IN REF. 1).  
 FT CONFLICT 129 131  
 FT MISSING (IN REF. 5).  
 FT CONFLICT 164 164  
 FT R -> P (IN REF. 3).  
 SQ SEQUENCE 270 AA; 29293 MW; B14BE0AAD3699AFE CRC64;  
 Query Match 36.2%; Score 513.5; DB 1; Length 270;  
 Best Local Similarity 41.4%; Pred. No. 4.7e-37;  
 Matches 113; Conservative 46; Mismatches 97; Indels 17; Gaps 10;  
 QY 2 LLSLTSLVLGSSWGCVPALTPTALSPALSNQIRIVNGENAVPGSWPQVSLQ-DNTG--FH 58  
 DB 1 MMLRLSSLLAVASGYPSPSRP-----SSRVYNGEDAVPYSPWQVSLQYKSGSYH 56  
 QY 59 FCGSLISPNWVVTAAHCQVTPGRHFVGLGEYDRSSNAEPVQLSI--ARAITHPNWNA 116  
 DB 57 TCGSLIADPWVVTAGHCISSSWTYQVGLGEYDRAVKEGEPQVPIPSNGDLFVHPLNRS 116  
 QY 117 TM--NNDLTLLKLSPARYTAQVSPVCLASTNEALPSGLTCVTGWRISGVGNVTPTARL 174  
 DB 117 CVACGNIDIALKLSRAQLGDAVLASLPAGDILNETPCYITGWRLYTNGPL-PDKL 175  
 QY 175 QQVPLPLTVNOCRQ--YWGARTIDAMICAGSGSCQDGGPLVC-QKGNFWLIGI 231  
 DB 176 QEALLPVDYEHCSRWNWNGSSVYKTMVCAGDIRSGCNGDGGPLNCPTEGGWQVHG 235  
 QY 232 VSW-GPKNCNI-QAPAMYTRVSKFSTWVQVMA 262  
 DB 236 TSEVSAFGCNTRKPKVFTVRSFAFDWIBETIA 268  
 RESULT 14  
 CAC3\_BOVIN  
 ID CAC3\_BOVIN STANDARD; PRT; 253 AA.  
 AC P05805;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Proproteinase E precursor (Procarboxypeptidase A complex component  
 DE III) (Procarboxypeptidase A-S6 subunit III) (PROCPA-S6 III).  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE OF 1-25.  
 RX MEDLINE=91099520; PubMed=2269366;  
 RA Pascual R., Vendrell J., Aviles F.X., Bonicel J., Wicker C.,



Search completed: December 20, 2002, 15:17:13  
Job time : 10.0435 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:14:28 : Search time 16.6957 Seconds  
(without alignments)  
1520.126 Million cell updates/sec

Title: US-09-856-319b-4

Perfect score: 1418

Sequence: 1 MLLSLTLTLVLLGSGGCG.....AMVTRVSKFTWVQVYAYN 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_73.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1269	89.5	264	2 I38136	chymotrypsin-like
2	832.5	58.7	263	2 A21195	chymotrypsin (EC 3
3	822.5	58.0	263	2 A31299	chymotrypsin (EC 3
4	798.5	56.3	263	2 S47537	chymotrypsin (EC 3
5	786.5	55.5	245	1 KYBOB	chymotrypsin (EC 3
6	781.5	55.1	263	1 KYRTB	chymotrypsin (EC 3
7	774.5	54.6	245	1 KYBOA	chymotrypsin (EC 3
8	737	52.0	244	2 S72219	chymotrypsin B - A
9	537	37.9	269	2 A26823	pancreatic elastase
10	517	36.5	271	1 ELRT2	pancreatic elastase
11	515	36.3	269	2 B26823	pancreatic elastase
12	513.5	36.2	270	2 B29934	pancreatic elastase
13	499	35.2	1524	2 T30337	polyprotein - Afri
14	495.5	34.9	268	2 S68826	pancreatic elastase
15	495.5	34.9	268	2 S68825	pancreatic elastase
16	492	34.7	271	2 A25528	pancreatic elastase
17	485.5	34.2	240	1 CPBOA3	procarboxypeptidase
18	485.5	34.2	270	2 A29934	pancreatic elastase
19	485.5	34.2	812	1 PLMS	plasmin (EC 3.4.21
20	484	34.1	810	1 PLHU	plasmin (EC 3.4.21
21	479	33.8	269	2 C26823	pancreatic elastase
22	477	33.6	126	2 A23473	chymotrypsin-like
23	477	33.6	460	2 B61545	plasmin (EC 3.4.21
24	475	33.5	810	2 B30848	plasmin (EC 3.4.21
25	474	33.4	790	1 PLPG	plasmin (EC 3.4.21
26	465	32.8	270	2 S56160	mast cell tryptase
27	463.5	32.7	273	2 A47246	tryptase (EC 3.4.2
28	462.5	32.6	343	1 A57014	proctasin (EC 3.4.
29	459	32.4	455	2 A61545	plasmin (EC 3.4.21

## ALIGNMENTS

### RESULT 1

I38136

chymotrypsin-like proteinase (EC 3.4.21.-) CTRL-1 - human

C:Species: Homo sapiens (man)

C:Date: 06-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change 22-Jun-1999

C:Accession: I38136

R:Larsen, F.; Solheim, J.; Kristensen, T.; Kolsto, A.B.; Prydz, H.

Hum. Mol. Genet. 2, 1589-1595, 1993

A:Title: A tight cluster of five unrelated human genes on chromosome 16q22.1.

A:Reference number: I38135; MUID:94093544; PMID:8268911

A:Accession: I38136

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-264 <RES>

A:Cross-references: EMBL:X71874; NID:g406226; PIDN:CAA50710.1; PID:g406228

C:Genetics:

A:Gene: GDB:CTRL

A:Cross-references: GDB:204061

A:Map position: 16q22.1-16q22.1

A:Introns: 18/1; 52/3; 79/2; 106/3; 167/1; 211/3

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; serine proteinase

F:34-257/Domain: trypsin homology <TRY>

F:75,121,214/Active site: His, Asp, Ser #status predicted

Query Match 89.5%; Score 1269; DB 2; Length 264;

Best Local Similarity 86.4%; Pred. No. 9.9e-98;

Matches 228; Conservative 25; Mismatches 11; Indels 0; Gaps 0;

Qy 1 MLLSLTLTLVLLGSGGCGVPAITPALSYNQIRVNGENAVPGSWPQVSLQDNTGFHFC 60

Db 1 MLLSLTLTLVLLGSGGCGIPAIKPAUSFSQIRVNGENAVLGSPWPQVSLQDSGGFHF 60

Qy 61 GGSLSIPNVVYTAACQVTPGRHFVVLGEYDRSSNAEPQVLSARAITHPNWNANTMNN 120

Db 61 GGSLSISQSVVYTAACNVSPGRHFVVLGEYDRSSNAEPQLQVLSVSRATHTPSWNTMNN 120

Qy 121 DLTLLKASPARYAQVSPVCLASTNEALPSGLTCVTTGWGRISGVGNVTPARLQOVVLP 180

Db 121 DVTLLKLASPAQYTRISPVCLASNEALTEGLTCVTTGWGRISGVGNVTPARLQOVVLP 180

Qy 181 LVTYNQCRQYNGARITDAMICAGSGASSCGDGGPLVCQKGNVTWVLIGVSNKTKCN 240

Db 181 LVTYNQCRQYNGSITDSMICAGGAGASSCGDGGPLVCQKGNVTWVLIGVSNKTKCN 240

Qy 241 IQAPAMVTRVSKFTWVQVYAYN 264

Db 241 VRAPAVTRVSKFTWVQVYAYN 264

### RESULT 2

A21195

1 MAFLWLLSCWALLGTTFCGCVPAIHPVLUSGLSRVINGEDAVPGSWPQVSLQDTGTFHFC 60

61 GGSLLSPNWNVYTAACHQVTPGRHFHVILGEYDRSSNAEPQVLSIARAIITHPNWNANTMNN 120

61 GGSLLSEDWVYTAACHGVRTS-DVVVAGEFDGSDDEENIQVLKIAKVPKPKFSILTIVNN 119

121 DUTLLKSLASPYRTAQSVPCLASTNEALPSGLTCVTVTGWGRISGVGNVTPARLQOVVLP 180

120 DITLLKLATPARFSQTVSAVCLPSADDDFPAGTLCATTGWGKTQYNANKTPDKLQQAALP 179

181 LVTVNQCQRYWCARTITDAMICAGGSGASSCGDGGPLVCQKQNTWTLVLIGVSGTKNCN 240

180 LLSNAECKSKSGRRITDYMICAGASGVSSCMGDSGGPLVCQKQGAWTLVGVISWGSOTCS 239

241 IQAPAMYTRVSKFSTWVINOVMAYN 264

240 TSSPGVYARVTKLIPWOKILAAAN 263

RESULT 4

S47537

chymotrypsin (EC 3.4.21.1) precursor - Atlantic cod

C:Species: Gadus morhua (Atlantic cod)

C:Date: 26-Dec-1994 #sequence\_revision 03-Aug-1995 #text\_change 22-Jun-1999

C:Accession: S47537; S43163

R:Gudmundsdottir, A.; Oskarsson, S.; Eakin, A.E.; Craik, C.S.; Bjarnason, J.B.

Biochim. Biophys. Acta 1219, 211-214, 1994

A:Title: Atlantic cod cDNA encoding a psychrophilic chymotrypsinogen.

A:Reference number: S47537; MUID:94368860; PMID:8086467

A:Accession: S47537

A:Molecule type: mRNA

A:Residues: 1-263 <GUD>

A:Cross-references: EMBL:X78490; NID:q468750; PIDN:CAA55242.1; PID:q468751

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; protein digestion; serine proteinase

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-263/Product: chymotrypsin #status predicted <MAT>

F:34-256/Domain: trypsin homology <TRY>

F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 56.3%; Score 798.5; DB 2; Length 263;

Best Local Similarity 56.9%; Pred. No. 8.8e-59;

Matches 144; Conservative 38; Mismatches 70; Indels 1; Gaps 1;

12 LIGSSWGCQVPAITPAISYNORIVNGENAVPGSPWQVSLQDNTGFHFCGGSLLSPNWNV 71

12 LFRRTYGGCRPAISVITGYSRIVNGEAVPHSWQVSLQDQTFHFCGGSLLINENNVV 71

72 TAAHCQVTPGRHFHVILGEYDRSSNAEPQVLSIARAIITHPNWNANTMNDLTLKLASPA 131

72 TAAHCNV-KNTHRVVLGENDRSNSEGQVQVMTVGQVFKHPRYNGFTINNDILLVKLATPA 130

132 RTAQSVPCLASTNEALPSGLTCVTTCWGRISGVGNVTPARLQOVVLPVLTVNQCQRYW 191

131 TLNMRVSPCLAEITDDVFEFGMKCVKTSQWGLTRYNAADTPALLOQAALPLLTNPQCKKFW 190

192 GARITDAMICAGGSGASSCGDGGPLVCQKQNTWTLVLIGVSGTKNCNIQAPAMYTRVS 251

191 GNKISDLMICAGAGASCMGDSGGPLVCQKQAGSWTLVGVISWGSOTCTPTMPGYIARVT 250

252 KFTSWVINOVMAYN 264

251 ELRAWVDQTIAAN 263

RESULT 5

KYBOB

chymotrypsin (EC 3.4.21.1) B precursor - bovine

N:Alternate names: chymotrypsinogen B

C:Species: Bos primigenius taurus (cattle)

C:Date: 08-Oct-1981 #sequence\_revision 08-Oct-1981 #text\_change 18-Jul-1997

R:Smillie, L.B.; Furka, A.; Nagabhushan, N.; Stevenson, K.J.; Parkes, C.O.

Nature 218, 343-346, 1968  
A:Title: Structure of chymotrypsinogen B compared with chymotrypsinogen A and trypsinogen  
A:Reference number: A00953; MUID:68238908; PMID:5649671  
A:Accession: A00953  
A:Molecule type: protein  
A:Residues: 1-245 <SM>  
C:Comment: Chymotrypsinogen B is synthesized, along with chymotrypsinogen A, in the acinar cells of the pancreas. The first activation cleavage, leading to pi-chymotrypsin B, occurs in the secretory granules.  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen  
F:1-15/Domain: propeptide #status experimental <PRO>  
F:16-245/Product: chymotrypsin B #status experimental <MAT>  
F:16-238/Domain: trypsin homology <TRY>  
F:1-122,42-58,136-201,168-182,191-220/Disulfide bonds: #status experimental  
F:57,102,195/Active site: His, Asp, Ser #status experimental

Query Match 55.5%; Score 786.5; DB 1; Length 245;  
Best Local Similarity 57.7%; Pred. No. 8e-58;  
Matches 142; Conservative 32; Mismatches 71; Indels 1; Gaps 1;

Qy 19 CGVPAITPALSIVRIVNGENAVPGSWPQVSLQDNTGTFHFCGGLISPNWVVTAAHCOV 78  
Db 1 CGVPAIQPVLSGLARIVNGEDAVPGSWPQVSLQDSTGFHFCGGLISEDWVVTAAHCOV 60

Qy 79 TPGRHFVVLGYDRSSNAEPVQVLSIARAIHPNNANTMNDLTLKLASPARYTAQVS 138  
Db 61 TTS-DVVVAGEFDQLEATEDQVLKIGKVFKNPKFSILTNRNDITLLKLATPAQFSETVS 119

Qy 139 PVCLASTNEALPSGLTCVTGHRISGVGNVTPARLQVVLPLVTVNQCQRYWAGARITDA 198  
Db 120 AVCLPSADEDFPAGMLCATGTGKTKYNALKTPDLQLOATLPIVNTDCRYKWSGRVTDV 179

Qy 199 MICAGSGASSCGSGGGLVCKQNTWVLIGIVSWGKNCNIQAPAMTVRSKFTSWIN 258  
Db 180 MICAGASGVSSCMGSGGLVCKQNGAWTLAGIVSWGSGTCTSTPVAIVARTALMPWQV 239

Qy 259 QVMAYN 264  
Db 240 ETLAN 245

RESULT 6  
KYTB  
chymotrypsin (EC 3.4.21.1) B precursor - rat  
N:Alternate names: chymotrypsinogen B  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 18-Jun-1999  
C:Accession: A22658  
R:Beil, G.I.; Quinto, C.; Quiroga, M.; Valenzuela, P.; Craik, C.S.; Rutter, W.J.  
J. Biol. Chem. 259, 14265-14270, 1984  
A:Title: Isolation and sequence of a rat chymotrypsin B gene.  
A:Reference number: A22658; MUID:85054881; PMID:6209274  
A:Accession: A22658  
A:Molecule type: DNA  
A:Residues: 1-263 <BE>  
A:Cross-references: GB:K02298; NID:g203653; PIDN:AAA98732.1; PID:g203654  
C:Genetics:  
A:Introns: 18/1; 52/3; 79/2; 105/3; 166/1; 210/3  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-33/Domain: propeptide #status predicted <PRO>  
F:34-263/Product: chymotrypsin B #status predicted <MAT>  
F:34-256/Domain: trypsin homology <TRY>  
F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 55.1%; Score 781.5; DB 1; Length 263;  
Best Local Similarity 53.8%; Pred. No. 2.2e-57;  
Matches 142; Conservative 40; Mismatches 81; Indels 1; Gaps 1;

Qy 1 MULLSLTSLVLGSGWCGVPAITPALSIVRIVNGENAVPGSWPQVSLQDNTGFHFC 60  
Db 1 MAPLWLVSCFALVGATFGCGVPTIQVLTGLSRVNGEDAIPEGSPWQVSLQDKTGFHFC 60

Qy 61 GGSLSIPNNVVTAAHCOVTPGRHFVVLGYDRSSNAEPVQVLSIARAIHPNNANTMNN 120  
Db 61 GGSLSIEDWVVTAAHCGVKTS-DVVVAGEFDQSGDEENIQVLKIAQVFNKFNFTVRN 119

Qy 121 DUTLLKLASPARYTAQVSPVCLASTNEALPSGLTCVTGHRISGVGNVTPARLQVVLPL 180  
Db 120 DITLLKLATPAQFSETVSACLPNVDDDFPCTVCATTGKTKYNALKTPKLOQAALP 179

Qy 181 LTVNQCQRYWAGARITDAMICAGSGASSCGGGLVCKQNTWVLIGIVSWGKNCN 240  
Db 180 IVSEADCKKSGSKITDVTMTACASGVSSCMGSGGLVCKQKDGWVTLTLAGIVSGSGVCS 239

Qy 241 IQAPAMTVRSKFTSWINQVMAYN 264  
Db 240 TSTPAVYSRVTPALMPWQVILEAN 263

## RESULT 7

## KYBOA

chymotrypsin (EC 3.4.21.1) A precursor - bovine

N:Alternate names: chymotrypsinogen A

C:Species: Bos primigenius taurus (cattle)

C:Date: 07-May-1981 #sequence\_revision 07-May-1981 #text\_change 07-May-1999

C:Accession: A30235; A93158; S29650; A00952

R:Brown, J.R.; Hartley, B.S.

Biochem. J. 101, 214-228, 1966

A:Title: Location of disulphide bridges by diagonal paper electrophoresis. The disulphide

A:Reference number: A90235; MUID:67181721; PMID:5971783

A:Accession: A90235

A:Molecule type: protein

A:Residues: 1-101,'N',103-245 <BRO>

R:Blow, D.M.; Birktoft, J.J.; Hartley, B.S.

Nature 221, 337-340, 1969

A:Title: Role of a buried acid group in the mechanism of action of chymotrypsin.

A:Reference number: A93158; MUID:69106266; PMID:5764436

A:Contents: annotation; revision to residue 102

R:Meloun, B.; Klueh, I.; Kostka, V.; Moravek, L.; Prusik, Z.; Vanacek, J.; Keil, B.; S.

Biochim. Biophys. Acta 130, 543-546, 1966

A:Title: Covalent structure of bovine chymotrypsinogen A.

A:Reference number: A90572; MUID:67183948; PMID:5972866

A:Accession: A93158

A:Molecule type: protein

A:Residues: 1-101,'N',103-245 <MEL>

A:Note: disulfide bonds were determined

R:Cutruzzola, F.; Ascenzi, P.; Barra, D.; Bolognesi, M.; Menegatti, E.; Sarti, P.; Sc

Biochim. Biophys. Acta 1161, 201-208, 1993

A:Title: Selective oxidation of Met-192 in bovine alpha-chymotrypsin. Effect on catal

A:Reference number: S29650; MUID:93160238; PMID:8431470

A:Accession: S29650

A:Molecule type: protein

A:Residues: 1-12;16-27;149-160;181-200 <CUT>

R:Smillie, L.B.; Hartley, B.S.

Biochem. J. 101, 232-241, 1966

A:Title: Histidine sequences in the active centres of some 'serine' proteinases.

A:Reference number: A90236; MUID:67181723; PMID:5971785

A:Contents: annotation; active site

R:Birktoft, J.J.; Blow, D.M.; Henderson, R.; Steitz, T.A.

Philos. Trans. R. Soc. Lond. B257, 67-76, 1970

A:Title: The structure of alpha-chymotrypsin.

A:Reference number: A93754

C:Contents: annotation; X-ray crystallography

C:Comment: Chymotrypsinogens are synthesized in the acinar cells of pancreas.

C:Comment: Tryptic cleavage after Arg-15 results in a fully active enzyme (pi-chymotri

delta-chymotrypsin; further chymotryptic cleavage liberates the dipeptide Thr-147 and

d Asn-148 directly from chymotrypsinogen, which leads to the degraded form neochymotr

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen

F:1-245/Product: chymotrypsinogen #status experimental <ZYM>

F:1-13,16-146,149-245/Product: alpha-chymotrypsin #status experimental <MPT>

F:16-238/Domain: trypsin homology <TRY>

F:1-122,42-58,136-201,168-182,191-220/Disulfide bonds: #status experimental

F:57,102,195/Active site: His, Asp, Ser #status experimental

Query Match 54.6%; Score 774.5; DB 1; Length 245;  
 Best Local Similarity 57.3%; Pred. No. 7.9e-57;  
 Matches 141; Conservative 32; Mismatches 72; Indels 1; Gaps 1;

QY 19 CGVPAITPALSINORIVNGENAVPGSWPQVSLQDNTGTFHFCGSLSPNNVYTAACHQV 78  
 Db 1 CGVPAIQVLSGLSRVINGEEAVPGSWPQVSLQDNTGTFHFCGSLSPNNVYTAACHQV 60

QY 79 TPGRHFFVLGEYDRSSNAEPVQVLSIARATHPHNWNANTWNNDLTLKLPASARYTAQVS 138  
 Db 61 TTS-DVVYAGEFDQSSSERIQKIAKVFKNKYNLSLTINDITLKLSTAAFSQSOTVS 119

QY 139 PVCLASTNEALPSGLTCTVTGWRISGVNGVTPARLQOVVPLPLVTNNOCQRYNGARITDA 198  
 Db 120 AVCLPSASDFAGTCTVTGWLTRYTNANTPDRLQOASLPDLSNTNCKRYGWTIKDA 179

QY 199 MICAGSGASSCGDGGPLVCQKGNVTWLVIGVSWGTKNCNIQAPAMYTRVSKFSTWIN 258  
 Db 180 MICAGASGVSCMGDGGPLVCCKNGAWTLVIGVSWGSSSTCTSTPGVYARVATLVNMQ 239

QY 259 QVMAYN 264  
 Db 240 QTLAAN 245

RESULT 8  
 S72219  
 Chymotrypsin B - Atlantic cod (fragments)  
 C:Species: Gadus morhua (Atlantic cod)  
 C:Date: 19-Mar-1998 #sequence\_revision 17-Apr-1998 #text\_change 07-Aug-1998  
 C:Accession: S72219  
 R:Leth-Larsen, R.; Asgeirsson, B.; Thorolfsson, M.; Norregaard-Madsen, M.; Hojrup, P.  
 Biochim. Biophys. Acta 1297, 49-56, 1996  
 A:Title: Structure of chymotrypsin variant B from Atlantic cod, Gadus morhua.  
 A:Reference number: S72219; MUID:96439045; PMID:88411380  
 A:Accession: S72219  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-14;15-244 <LET>  
 C:Superfamily: trypsin; trypsin homology  
 F:15-237/Domain: trypsin homology <TRY>

Query Match 52.0%; Score 737; DB 2; Length 244;  
 Best Local Similarity 54.2%; Pred. No. 9.9e-54;  
 Matches 135; Conservative 43; Mismatches 63; Indels 8; Gaps 6;

QY 19 CGVPAITPALSINORIVNGENAVPGSWPQVSLQDNTGTFHFCGSLSPNNVYTAACHQV 78  
 Db 1 CGSPAIOQVPT-GYAIYNGEEAVPHSPWQVSLQDNTGTFHFCGSLSPNNVYTAACHQV 59

QY 79 TPGRHFFVLGEYDRSSNAEPVQVLSIARATHPHNWNANTWNNDLTLKLPASARYTAQVS 138  
 Db 60 RT-YHRVIVGEHDKASD-ENIQILKPSWTFPHKWDSTINNDISLIKLPASPAVLGTNVS 117

QY 139 PVCLASTNEALPSGLTCTVTGWRISGVNGVTPARLQOVVPLPLVTNNOCQRYNGAR-I 195  
 Db 118 PVCLGESSDFAPCMKCVTSWGGLTRYNAPG--TPNKLOQAALPLMSNECSQTWGNMI 175

QY 196 TDAMICAGSGASSCGDGGPLVCQKGNVTWLVIGVSWGTKNCNIQAPAMYTRVSKFST 255  
 Db 176 SDVMICAGAGATSCMGDGGPLVCQKDNVTWLVIGVSWGSSRCSTVTPAYARVTELRG 235

QY 256 WINOVAYN 264  
 Db 236 WVDILAN 244

RESULT 9  
 A26823  
 pancreatic elastase II (EC 3.4.21.71) precursor - pig  
 N:Alternate names: pancroateptidase E  
 C:Species: Sus scrofa domestica (domestic pig)

C:Date: 16-Aug-1988 #sequence\_revision 16-Aug-1988 #text\_change 22-Jun-1999  
 C:Accession: A26823  
 R:Kawahana, I.; Tani, T.; Shimoda, K.; Takiguchi, Y.  
 DNA 6, 163-172, 1987  
 A:Title: Characterization of pancreatic elastase II cDNAs: two elastase II mRNAs are  
 A:Reference number: A0958; MUID:87217962; PMID:3646943  
 A:Accession: A26823  
 A:Molecule type: mRNA  
 A:Residues: 1-269 <KAW>  
 A:Cross-references: GB:M16651; NID:gl64441; PIDN:AAA1027.1; PID:gl64442  
 C:Superfamily: trypsin; trypsin homology  
 C:Keywords: hydrolase; serine proteinase  
 F:1-16/Domain: signal sequence #status predicted <SIG>  
 F:17-28/Domain: propeptide #status predicted <PRO>  
 F:29-269/Product: elastase II #status predicted <MAT>  
 F:29-262/Domain: trypsin homology <TRY>  
 F:73,121,216/Active site: His, Asp, Ser #status predicted

Query Match 37.9%; Score 537; DB 2; Length 269;  
 Best Local Similarity 43.1%; Pred. No. 4e-37;  
 Matches 118; Conservative 48; Mismatches 92; Indels 16; Gaps 11;

QY 2 LLLSLTSLVLLGSSWCGVPAITPALSINORIVNGENAVPGSWPQVSLQ-DNTG--FH 58  
 Db 1 MIRALLSTLVAG-ALSCGLPANLPQL---PRVVGEDARPNSWPQVSLQYDSSGQWRH 56

QY 59 FCGSLTSPNNVYTAACHQVTPGRHFVVLGEYDRSSNAEPVQVLSIARATHPHNWNANTM 118  
 Db 57 TCGTLVDQSWVLAHCISSTRTYRVLGRSLSTNEPGSLAVKVKLVVHQNNSQL 116

QY 119 N--NDITLLKLSPARYTAQVSPVCLASTNEALPSGLTCTVTGWRISGVNGVTPARLQ 176  
 Db 117 SGNNDIALLKLSPVSLTDRQLGLCPAAGTILPNKYCVYVTCWGRQLTNG-ASPDILQ 175

QY 177 VVPLVTVNOCQ--YNGARITDAMICAGSG-ASSCOGDSGGLVLCQKN-TWVLIGIV 232  
 Db 176 GQLLVVDYATCPKPGWGWSTVTKNMICAGDGISSCGDGGPLNCGGANGQMVHGIV 235

QY 233 SWGTK-NCN-IOAPAMYTRVSKFSTWINQVMAYN 264  
 Db 236 SFGSSLCGNYHKPSVFTRVSNVIDWINSVIANN 269

RESULT 10  
 ELRT2  
 pancreatic elastase II (EC 3.4.21.71) precursor - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 18-Aug-1982 #sequence\_revision 18-Aug-1982 #text\_change 24-Sep-1999  
 C:Accession: A00961  
 R:MacDonald, R.J.; Swift, G.H.; Quinto, C.; Swain, W.; Pictet, R.L.; Nikovits, W.; Ru  
 Biochemistry 21, 1453-1463, 1982  
 A:Title: Primary structure of two distinct rat pancreatic preproelastases determined  
 A:Reference number: A00960; MUID:82182967; PMID:6918221  
 A:Accession: A00961  
 A:Molecule type: mRNA  
 A:Residues: 1-271 <NAC>  
 A:Cross-references: GB:I00124; GB:J00731; NID:g204019; PIDN:AAA98780.1; PID:g204021  
 C:Superfamily: trypsin; trypsin homology  
 C:Keywords: hydrolase; pancreas; serine proteinase; zymogen  
 F:1-16/Domain: signal sequence #status predicted <SIG>  
 F:17-30/Domain: activation peptide #status predicted <ACT>  
 F:31-271/Product: elastase II #status predicted <MPT>  
 F:73-264/Domain: trypsin homology <TRY>  
 F:75,123,218/Active site: His, Asp, Ser #status predicted

Query Match 36.5%; Score 517; DB 1; Length 271;  
 Best Local Similarity 41.1%; Pred. No. 1.8e-35;  
 Matches 113; Conservative 46; Mismatches 100; Indels 16; Gaps 11;

QY 2 LLLSLTSLVLLGSSWCGVPAITPALSIN-ORIVNGENAVPGSWPQVSLQ---DNTGF 57  
 Db 1 MIRTLLLSLVAG-ALSCGYP--TYEQVHDSRVVGQEAAPNSWPQVSLQYLSGSKWH 57



A:Accession: S04490  
A:Molecule type: protein  
A:Residues: 94-128;132-164 <WEN>  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: glycoprotein; hydrolase; serine proteinase; zymogen  
F:1/7/Domain: signal sequence #status predicted <SIG>  
F:18-28/Domain: activation peptide #status predicted <ACT>  
F:29-270/Product: (or 31-270) pancreatic elastase IIB #status predicted <MAT>  
F:29-263/Domain: trypsin homology <TRY>  
F:73,123,217/Active site: His, Asp, Ser #status predicted  
F:114/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F:153/Binding site: carbohydrate (Asn) (covalent) #status absent

Query Match 36.2%; Score 513.5; DB 2; Length 270; ...  
Best Local Similarity 41.4%; Pred. No. 3.5e-35;  
Matches 113; Conservative 46; Mismatches 97; Indels 17; Gaps 10

QY 2 LLLSLTLVLGLSGGCGVPATIPALSYNQRIYNGENAVPGSWPQVSLQ-DWTG--FH 58  
DB 1 MMLRLSSLLLVAVASGYPGPPSRP-----SSXVNGEDAVPYSWPQVSLQYKSGSFYH 56  
QY 59 FCGSLISPNVYVTAACHCVTPGRHFVVLGEYDRSSNAEPQVQLSI--ARAITHPNNAN 116  
DB 57 TCGSLIAPDWWYTAGHCISSTRTQVVLGEYDRAVKEGPEQVPIPSGDLFVHPLNRS 116  
QY 117 TM--NNDLTLLKLSPARTYAQVSPVCLASTNEALPSGLTCVTTWGKRISGVGNVTPARL 174  
DB 117 CVACGNDIALIKLSRAQLGDVAQLASLPAGDILPNETPCVITGWGRLYFNGPL-PDKL 175  
QY 175 QQVVLPLVTYNQCRQ--YNGARTDAMICAGSGGASSCGSGGGLPVC-QKGNWTWVLIGI 231  
DB 176 QEALLPVVDYEHCSRNNWGWSSVKYTMWCAGDIIIRGNCNGSGGGLNCPTEDDGQWQHV 235  
QY 232 VSW-GTKNCNI-QAPAMYTRVSKFTWVNOVMA 262  
DB 236 TSFVSAGCNRTRKPVFTVRSAFIDWIEETIA 268

RESULT 13  
T30337  
Polyprotein - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 22-Oct-1999  
C:Accession: T30337  
R:Yang, J.C.; Lindsay, L.L.; Hedrick, J.L.  
submitted to the EMBL Data Library, March 1998  
A:Description: cDNA cloning of ovocytinase, a chymotrypsin-like protease released  
A:Reference number: Z20829  
A:Accession: T30337  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1524 <YAN>  
A:Cross-references: EMBL:U01290; NID:g2981640; PID:g2981641; PIDN:AAC24717.1

Query Match 35.2%; Score 499; DB 2; Length 1524;  
Best Local Similarity 36.8%; Pred. No. 4e-33;  
Matches 99; Conservative 52; Mismatches 82; Indels 36; Gaps 7

QY 19 CGVPAITPALSY-----NQRIYNGENAVPGSWPQVSLQDNTGFHRCGG 62  
DB 1264 CGNTVDPNMLYMARSGKIRNLNKGKSGRYVGGQAAPRSWPVLVSIQNSKRHYCGG 1323  
QY 63 SLISPNVYVTAACHCVTPGRHFVVLGEYDRS--SNAEPQVQLSIARAITHPNNANTMNN 120  
DB 1324 IIITNKKLLTAACEVKINLHRVVGGTDLTEVQNEHAKVYNSHVHELYMP--GSSPPRN 1381  
QY 121 DLTLKLSPARTYAQVSPVCLASTNEALPSGLT-----CVTTWGRISGVGNVT-PA 172  
DB 1382 DLLLELDTPLLLNSVAVIC-----LPDDVTDWTQAECLVAGWG-VTDVGMSLPT 1433  
QY 173 RLQOVVPLVTVNQRYGARITDAMICAGSGGASSCGSGGGLPVCOKGNWTWVLIGIV 232  
DB 1434 KLOAKVPIVSTKCKDYSDVTDNNICAGKAGASSCCMGDSGGPLICKRDRYLVGV 1493

F:30-268/Product: pancreatic elastase isoform 1 #status predicted <MAT>  
F:30-262/Domain: trypsin homology <TRY>

Query Match 34.98; Score 495.5; DB 2; Length 268;  
Best Local Similarity 40.1%; Pred. No. 1.1e-33;  
Matches 108; Conservative 46; Mismatches 102; Indels 13; Gaps 9;

Qy	3	LLSLTSLVLLGSSWGCGVPAITPALSYNORIVNGENAVPGSWPWOVSLQ--DNTGPHF	59
Db	1	MLGITVLAALLACASSCGVPSPFNLS--ARVGGEDARPHSWPQISLQYLKNDTWRT	58
Qy	60	CGSLISPNWYVTAACHCOVTPGRHFVVLGEYD-RSSNAEPVQVLSIARAI THPNMNA	118
Db	59	CGGTLIASNEFVLTAAHCISNRTYRVAVGKKNLEVEDEEGSLFVGVDTHVHKRNALL	118
Qy	119	NNDLTLLKLSPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQOV	178
Db	119	RNDIALIKLAHVELSDTIQVACLPEKDSLPLKDYPCYVTGWRLTWNGPIAD-KLQOGL	177
Qy	179	LPLVTVNOCRO--YVGARITDAMICAGGSA-SSCGDSGGPLVCOKGN-TWVLIGIVSW	234
Db	178	QPVDHATCSRIDWNGFVKKTMVCAGGDGVISACNGDSGGPLNCQLENGSWEVFGIVSF	237
Qy	235	GT-KNCNI-QAPAMYTRVSKFSTWINQVM	261
Db	238	GSRRCNTRKKPVVYTRVSAYIDWINEKM	266

Search completed: December 20, 2002, 15:19:48  
Job time : 17.6957 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:14:28 ; Search time 14.6087 Seconds  
(without alignments)  
1520.126 Million cell updates/sec

Title: US-09-856-319b-2\_COPY\_1\_231

Perfect score: 1221

Sequence: 1 MLLSLTLVLVLLGSSGGG.....GDSGGLVLCQKGTWVLIGI 231

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1214	99.4	264	2	I38136 chymotrypsin-like
2	714.5	58.5	263	2	A21195 chymotrypsin (EC 3.4.21.1)
3	701.5	57.5	263	2	S47537 chymotrypsin (EC 3.4.21.1)
4	689.5	56.5	263	2	A31299 chymotrypsin (EC 3.4.21.1)
5	670.5	54.9	263	1	KYRTB chymotrypsin (EC 3.4.21.1)
6	666.5	54.6	245	1	KYBOB chymotrypsin (EC 3.4.21.1)
7	657.5	53.8	245	1	KYBOA chymotrypsin (EC 3.4.21.1)
8	653	53.5	244	2	S72219 chymotrypsin B - A
9	472	38.7	269	2	A26823 pancreatic elastase
10	453.5	37.1	270	2	B29934 pancreatic elastase
11	451	36.9	126	2	A23473 chymotrypsin-like
12	445	36.4	269	2	B26823 pancreatic elastase
13	438	35.9	271	1	ELRT2 pancreatic elastase
14	427	35.0	269	2	C26823 pancreatic elastase
15	420.5	34.4	270	2	A29934 pancreatic elastase
16	415	34.0	271	2	A25528 pancreatic elastase
17	410.5	33.6	343	1	A57014 proctasin (EC 3.4.21.1)
18	409	33.5	1524	2	T30337 polyprotein - Afri
19	404	33.1	460	2	B61545 plasmin (EC 3.4.21.1)
20	397	32.5	810	1	PLHU plasmin (EC 3.4.21.1)
21	396	32.4	810	2	B30848 plasmin (EC 3.4.21.1)
22	395.5	32.4	1035	1	A43090 enteropeptidase (E
23	391	32.0	790	1	PLPG plasmin (EC 3.4.21.1)
24	390.5	32.0	268	2	S68826 pancreatic elastase
25	390.5	32.0	268	2	S68825 pancreatic elastase
26	389.5	31.9	240	1	CPBOA3 procaryopeptidase
27	389.5	31.9	812	1	PLMS plasmin (EC 3.4.21.1)
28	389	31.9	455	2	A61545 plasmin (EC 3.4.21.1)
29	388	31.8	270	2	S56160 mast cell tryptase

30	386	31.6	786	1	A47547 serine proteinase
31	386	31.6	1004	2	T30338 oviductin (EC 3.4.21.1)
32	385	31.5	248	2	S55067 trypsin (EC 3.4.21.1)
33	384.5	31.5	238	2	S31779 trypsin (EC 3.4.21.1)
34	384.5	31.5	273	2	A47246 trypsin (EC 3.4.21.1)
35	383.5	31.4	246	1	TRRT1 trypsin (EC 3.4.21.1)
36	383.5	31.4	246	1	TRDGC trypsin (EC 3.4.21.1)
37	382	31.3	1019	1	A56318 enteropeptidase (E
38	379	31.0	248	2	S55066 trypsin (EC 3.4.21.1)
39	378.5	31.0	638	1	KQHUP plasma kallikrein
40	377.5	30.9	812	1	PLBO plasmin (EC 3.4.21.1)
41	377.5	30.9	1113	2	JEO315 low-density lipopr
42	377	30.9	231	1	TRPCTR trypsin (EC 3.4.21.1)
43	376.5	30.8	268	2	J01473 pancreatic elastase
44	376	30.8	266	1	ELRT1 pancreatic elastase
45	375.5	30.8	247	1	TRDG trypsin (EC 3.4.21.1)

## ALIGNMENTS

### RESULT 1

I38136  
chymotrypsin-like proteinase (EC 3.4.21.-) CTRL-1 - human  
C:Species: Homo sapiens (man)  
C:Date: 06-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change 22-Jun-1999  
C:Accession: I38136  
R:Larsen, F.; Solheim, J.; Kristensen, T.; Kolsto, A.B.; Prydz, H.  
Hum. Mol. Genet. 2, 1589-1595, 1993  
A:Title: A tight cluster of five unrelated human genes on chromosome 16q22.1.  
A:Reference number: I38135; MUID:94093544; PMID:8268911  
A:Accession: I38136  
A:Status: Preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-264 <RES>  
A:Cross-references: EMBL:X71874; NID:g406226; PIDN:CAA50710.1; PID:g406228  
C:Genetics:  
A:Gene: GDB:CTRL  
A:Cross-references: GDB:204061  
A:Map position: 16q22.1-16q22.1  
A:Introns: 18/1; 52/3; 79/2; 106/3; 167/1; 211/3  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; serine proteinase  
F:34-257/Domain: trypsin homology <TRY>  
F:75.121.214/Active site: His, Asp, Ser #status predicted

Query Match 99.4%; Score 1214; DB 2; Length 264;  
Best Local Similarity 99.6%; Pred. No. 3.2e-95;  
Matches 230; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	MLLSLTSLVLLGSSGCGIPAIKPAISQRIVNGENAVLGSWPQVSLQDSSGFHFC	60
Db	1	MLLSLTSLVLLGSSGCGIPAIKPAISQRIVNGENAVLGSWPQVSLQDSSGFHFC	60
QY	61	GGSLISQSVVTAACHNCVSPGRHFVYVGLGYDRSSNAEPLQVLSVRAITHPSWNSITMNN	120
Db	61	GGSLISQSVVTAACHNCVSPGRHFVYVGLGYDRSSNAEPLQVLSVRAITHPSWNSITMNN	120
QY	121	DVTLKLAAPQYTTTRISPVCLASNEALTEGLTCVTGWRGLSGVGNVTPAHLQVALP	180
Db	121	DVTLKLAAPQYTTTRISPVCLASNEALTEGLTCVTGWRGLSGVGNVTPAHLQVALP	180
QY	181	LVTVNOCROYWSSITDSMICAGGAGSCQDGGGLVCOKGNTWVLIGI	231
Db	181	LVTVNOCROYWSSITDSMICAGGAGSCQDGGGLVCOKGNTWVLIGI	231

### RESULT 2

A21195  
chymotrypsin (EC 3.4.21.1) 2 precursor - dog  
C:Species: Canis lupus familiaris (dog)  
C:Date: 27-Jul-1990 #sequence\_revision 27-Jul-1990 #text\_change 22-Jun-1999  
C:Accession: A21195

R:Pinsky, S.D.; LaForge, K.S.; Luc, V.; Scheele, G.  
Proc. Natl. Acad. Sci. U.S.A. 80, 7486-7490, 1983  
A:Title: Identification of cDNA clones encoding secretory isoenzyme forms: sequence data  
A:Reference number: A21195; MUID:84170253; PMID:6584866  
A:Accession: A21195  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-263 <PIN>  
A:Cross-references: GB:K01173; NID:g163945; PIDN:AAA30841.1; PID:g163946  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; protein digestion; serine proteinase  
F:34-256/Domain: trypsin homology <TRY>  
F:75,120,213/Active site: His, Asp, Ser #status predicted

*Indel*

Query Match 58.5%; Score 714.5; DB 2; Length 263;  
Best Local Similarity 56.7%; Pred. No. 5e-53;  
Matches 131; Conservative 35; Mismatches 64; Indels 1; Gaps 1;  
QY 1 MLLSLTSLVLGSGGCGIPALPALSFSQRIVNGENAVLGSWPQVSLQDSSGFHFC 60  
DB 1 MAFLLSLSCFALLGTAFCGCGVPAIQPVLSGLSRVNGEDAVPGSWPQVSLQDSTGFHFC 60  
QY 61 GGLISQSWVVTAAHCNVSPGRHFVYLGEYDRSSNAEPLOVLSVRAITHPSWNSTTMNN 120  
DB 61 GGLISQSWVVTAAHCNVSPGRHFVYLGEYDRSSNAEPLOVLSVRAITHPSWNSTTMNN 119  
QY 121 DVTLKLLASPAQVYTRISPVCLASNEALTEGLTCVTTGRLSGVGNVTPAHLQVQVALP 180  
DB 120 DVTLKLLATPAREFSQVSAVCLPQATDDFPAGTLCVTTGRLKHTNANTPDKLQQAALP 179  
QY 181 LVTVNOCROYWDSITDSMTICAGAGASSCGDGGPLVCQKGNVTWVLGI 231  
DB 180 LLSNAECKKFWGSKITDLMVCAGAGSVSCMGDGGPLVCQKGNVTWVLGI 230

RESULT 3  
S47537  
chymotrypsin (EC 3.4.21.1) precursor - Atlantic cod  
C:Species: Gadus morhua (Atlantic cod)  
C:Date: 26-Dec-1994 #sequence\_revision 03-Aug-1995 #text\_change 22-Jun-1999  
C:Accession: S47537; S43163  
R:Gudmundsdottir, A.; Oskarsson, S.; Eakin, A.E.; Craik, C.S.; Bjarnason, J.B.  
Biochim. Biophys. Acta 1219, 211-214, 1994  
A:Title: Atlantic cod cDNA encoding a psychrophilic chymotrypsinogen.  
A:Reference number: S47537; MUID:94368860; PMID:8086467  
A:Accession: S47537  
A:Molecule type: mRNA  
A:Residues: 1-263 <GUD>  
A:Cross-references: EMBL:X78490; NID:g468750; PIDN:CAA55242.1; PID:g468751  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; protein digestion; serine proteinase  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-263/Product: chymotrypsin #status predicted <PRO>  
F:34-256/Domain: trypsin homology <TRY>  
F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 57.5%; Score 701.5; DB 2; Length 263;  
Best Local Similarity 58.6%; Pred. No. 6.3e-52;  
Matches 129; Conservative 33; Mismatches 57; Indels 1; Gaps 1;  
QY 12 LLGSSGCGIPALPALSFSQRIVNGENAVLGSWPQVSLQDSSGFHFCGGLISQSWV 71  
DB 12 LFRRTGCGRPALSPVITGYSRVINGEAVPHSMQVSLQDQTFHFCGGLINENWV 71  
QY 72 TAAHCNVSPGRHFVYLGEYDRSSNAEPLOVLSVRAITHPSWNSTTMNNVTLKLLASPA 131  
DB 72 TAAHCNVK-NYHRAVLGEHDDSSNSEQVQVMTGVQFKHPRYNGFTINNDILLVKTATPA 130  
QY 132 QYTRISPVCLASNEALTEGLTCVTTGRLSGVGNVTPAHLQVQVALPVTNOCROYW 191  
DB 131 TLNMRVSPVCLATDDVFEGGCMKVTSGWGLTRYNADPTALLQQAALPLLTNEQCKKF 190  
QY 192 DSSITDSMTICAGAGASSCGDGGPLVCQKGNVTWVLGI 231

DB 191 GNKISDLMICAGAGASSCMGDSGGPLVCQKAGSWTLVGI 230  
RESULT 4  
A31299  
chymotrypsin (EC 3.4.21.1) precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 08-Jun-1989 #sequence\_revision 08-Jun-1989 #text\_change 22-Jun-1999  
C:Accession: A31299  
R:Tomita, N.; Izumoto, Y.; Horii, A.; Doi, S.; Yokouchi, H.; Ogawa, M.; Mori, T.; Mat  
Biochem. Biophys. Res. Commun. 158, 569-575, 1989  
A:Title: Molecular cloning and nucleotide sequence of human pancreatic prechymotrypsin  
A:Reference number: A31299; MUID:89134264; PMID:2917002  
A:Accession: A31299  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-263 <TOM>  
A:Cross-references: GB:M24400; NID:g181189; PIDN:AAA52128.1; PID:g181190  
C:Genetics:  
A:Gene: GDB:CTRB1; CTRB  
A:Cross-references: GDB:119820; OMIM:118890  
A:Map position: 16q23.1-16q23.1  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; protein digestion; serine proteinase  
F:34-256/Domain: trypsin homology <TRY>  
F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 56.5%; Score 689.5; DB 2; Length 263;  
Best Local Similarity 55.4%; Pred. No. 6.5e-51;  
Matches 128; Conservative 35; Mismatches 67; Indels 1; Gaps 1;  
QY 1 MLLSLTSLVLGSGGCGIPALPALSFSQRIVNGENAVLGSWPQVSLQDSSGFHFC 60  
DB 1 MAFLLSLSCWALLGTTFCGCGVPAIHPVLSGLSRVNGEDAVPGSWPQVSLQDSTGFHFC 60  
QY 61 GGLISQSWVVTAAHCNVSPGRHFVYLGEYDRSSNAEPLOVLSVRAITHPSWNSTTMNN 120  
DB 61 GGLISQSWVVTAAHCNVSPGRHFVYLGEYDRSSNAEPLOVLSVRAITHPSWNSTTMNN 119  
QY 121 DVTLKLLASPAQVYTRISPVCLASNEALTEGLTCVTTGRLSGVGNVTPAHLQVQVALP 180  
DB 120 DITLLLATPAREFSQVSAVCLPSADDDFPAGTLCATGKTKYNANKTPDKLQQAALP 179  
QY 181 LVTVNOCROYWDSITDSMTICAGAGASSCGDGGPLVCQKGNVTWVLGI 231  
DB 180 LLSNAECKKSWGRITDVMICAGAGSVSCMGDGGPLVCQKGNVTWVLGI 230

RESULT 5  
KYRTB  
chymotrypsin (EC 3.4.21.1) B precursor - rat  
N:Alternate names: chymotrypsinogen B  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 18-Jun-1999  
C:Accession: A22658  
R:Bell, G.I.; Quinto, C.; Quiroga, M.; Valenzuela, P.; Craik, C.S.; Rutter, W.J.  
J. Biol. Chem. 259, 14265-14270, 1984  
A:Title: Isolation and sequence of a rat chymotrypsin B gene.  
A:Reference number: A22658; MUID:85054881; PMID:6209274  
A:Accession: A22658  
A:Molecule type: DNA  
A:Residues: 1-263 <BEL>  
A:Cross-references: GB:K02298; NID:g203653; PIDN:AAA98732.1; PID:g203654  
C:Genetics:  
A:Introns: 18/1; 52/3; 79/2; 105/3; 166/1; 210/3  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-33/Domain: propeptide #status predicted <PRO>  
F:34-263/Product: chymotrypsin B #status predicted <MAT>  
F:34-256/Domain: trypsin homology <TRY>  
F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 56.5%; Score 689.5; DB 2; Length 263;  
Best Local Similarity 55.4%; Pred. No. 6.5e-51;  
Matches 128; Conservative 35; Mismatches 67; Indels 1; Gaps 1;  
QY 1 MLLSLTSLVLGSGGCGIPALPALSFSQRIVNGENAVLGSWPQVSLQDSSGFHFC 60  
DB 1 MAFLLSLSCWALLGTTFCGCGVPAIHPVLSGLSRVNGEDAVPGSWPQVSLQDSTGFHFC 60  
QY 61 GGLISQSWVVTAAHCNVSPGRHFVYLGEYDRSSNAEPLOVLSVRAITHPSWNSTTMNN 120  
DB 61 GGLISQSWVVTAAHCNVSPGRHFVYLGEYDRSSNAEPLOVLSVRAITHPSWNSTTMNN 119  
QY 121 DVTLKLLASPAQVYTRISPVCLASNEALTEGLTCVTTGRLSGVGNVTPAHLQVQVALP 180  
DB 120 DITLLLATPAREFSQVSAVCLPSADDDFPAGTLCATGKTKYNANKTPDKLQQAALP 179  
QY 181 LVTVNOCROYWDSITDSMTICAGAGASSCGDGGPLVCQKGNVTWVLGI 231  
DB 180 LLSNAECKKSWGRITDVMICAGAGSVSCMGDGGPLVCQKGNVTWVLGI 230

Query Match 54.98; Score 670.5; DB 1; Length 263;  
Best Local Similarity 53.28; Pred. No. 2.6e-49;  
Matches 123; Conservative 35; Mismatches 72; Indels 1; Gaps 1;  
Qy 1 MLLSLSLVLLGSSWCGGIPALPAKSFQRIVNGENAVLGSWPQVSLQDSSGFHFC 60  
Db 1 MAFWLWVSCFALVATGCGVPTIQLVLTGLSRVNGEDAIPTGSPWQVSLQDKTGPHFC 60  
Qy 61 GGSLSOSVWVTAHNCVNSPGRHFVGLGEYDRSSNAEPLQVLSVRAITHPSWNTMNN 120  
Db 61 GGSLSIEDWVWTAHNCVNSPGRHFVGLGEYDRSSNAEPLQVLSVRAITHPSWNTMNN 119  
Qy 121 DVTLLKLASPAQVYTRTSIPVCLASSNEALTEGLTCVTTGWLGRSLVGNVTFAHLQVVALP 180  
Db 120 DITLLKLATPAQFSETVSAVCLPNVDDPPETVCATITGKTKYNALKTEPEKIQQAALP 179  
Qy 181 LVTYNQCRQYWDSSITDSMICAGAGASSCGDGGPLVCOKGNTWVLIGI 231  
Db 180 IVSEADCKKSGSKITDVTMTAGASGVSSCMGDSGGPLVCQKDGWVTLAGI 230

## RESULT 6

KYOBO

chymotrypsin (EC 3.4.21.1) B precursor - bovine  
N:Alternate names: chymotrypsinogen B  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 08-Oct-1981 #sequence\_revision 08-Oct-1981 #text\_change 18-Jul-1997  
C:Accession: A00953  
Nature 218, 343-346, 1968  
A:Title: Structure of chymotrypsinogen B compared with chymotrypsinogen A and trypsinogen A  
A:Reference number: A00953; MUID:68238908; PMID:5649671  
A:Accession: A00953  
A:Molecule type: protein  
A:Residues: 1-245 <SMI>  
C:Comment: Chymotrypsinogen B is synthesized, along with chymotrypsinogen A, in the acinar cells of the pancreas.  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen  
F:1-15/Domain: propeptide #status experimental <PRO>  
F:16-245/Product: chymotrypsin B #status experimental <MAT>  
F:16-238/Domain: trypsin homology <TRY>  
F:1-122,42-58,136-201,168-182,191-220/Disulfide bonds: #status experimental  
F:57,102,195/Active site: His, Asp, Ser #status experimental

Query Match 54.68; Score 666.5; DB 1; Length 245;  
Best Local Similarity 57.38; Pred. No. 5.2e-49;  
Matches 122; Conservative 28; Mismatches 62; Indels 1; Gaps 1;  
Qy 19 CGIPALPAKSFQRIVNGENAVLGSWPQVSLQDSSGFHFCGSLISQSWVWVTAHNCV 78  
Db 1 CGVPAIQVLSGLARVNGEDAVPGSWPQVSLQDSTGFHFCGSLISEDWVWVTAHNCV 60  
Qy 79 SPGRHFVVLGEYDRSSNAEPLQVLSVRAITHPSWNTMNNVTLKLASPAQVYTRIS 138  
Db 61 TTS-DVVVAGEFDQGLETDQVLKIGKVFKNPFSILTVRNDITLLKATPAQFSETVS 119  
Qy 139 PVCLASSNEALTEGLTCVTTGWLGRSLVGNVTFAHLQVVALPVTYNOCROYWDSSTIDS 198  
Db 120 AVCLPSADEDPFAGMLCATTCGKTKYNALKTPKLOQAATLPVSNNTDCRYWGSRTVDV 179  
Qy 199 MICAGGAGASSCGDGGPLVCOKGNTWVLIGI 231  
Db 180 MICAGASVSSCMGDSGGPLVCQKNGAWTLAGI 212

## RESULT 7

KYBOA

chymotrypsin (EC 3.4.21.1) A precursor - bovine  
N:Alternate names: chymotrypsinogen A  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 07-May-1981 #sequence\_revision 07-May-1981 #text\_change 07-May-1999

C:Accession: A90235; A93158; S29650; A00952  
R:Brown, J.R.; Hartley, B.S.  
Biochem. J. 101, 214-228, 1966  
A:Title: Location of disulphide bridges by diagonal paper electrophoresis. The disulphide bridges are located at positions 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 841, 842, 843, 844, 845, 846, 847, 848, 849, 850, 851, 852, 853, 854, 855, 856, 857, 858, 859, 860, 861, 862, 863, 864, 865, 866, 867, 868, 869, 870, 871, 872, 873, 874, 875, 876, 877, 878, 879, 880, 881, 882, 883, 884, 885, 886, 887, 888, 889, 890, 891, 892, 893, 894, 895, 896, 897, 898, 899, 900, 901, 902, 903, 904, 905, 906, 907, 908, 909, 910, 911, 912, 913, 914, 915, 916, 917, 918, 919, 920, 921, 922, 923, 924, 925, 926, 927, 928, 929, 930, 931, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, 972, 973, 974, 975, 976, 977, 978, 979, 980, 981, 982, 983, 984, 985, 986, 987, 988, 989, 990, 991, 992, 993, 994, 995, 996, 997, 998, 999, 1000, 1001, 1002, 1003, 1004, 1005, 1006, 1007, 1008, 1009, 1010, 1011, 1012, 1013, 1014, 1015, 1016, 1017, 1018, 1019, 1020, 1021, 1022, 1023, 1024, 1025, 1026, 1027, 1028, 1029, 1030, 1031, 1032, 1033, 1034, 1035, 1036, 1037, 1038, 1039, 1040, 1041, 1042, 1043, 1044, 1045, 1046, 1047, 1048, 1049, 1050, 1051, 1052, 1053, 1054, 1055, 1056, 1057, 1058, 1059, 1060, 1061, 1062, 1063, 1064, 1065, 1066, 1067, 1068, 1069, 1070, 1071, 1072, 1073, 1074, 1075, 1076, 1077, 1078, 1079, 1080, 1081, 1082, 1083, 1084, 1085, 1086, 1087, 1088, 1089, 1090, 1091, 1092, 1093, 1094, 1095, 1096, 1097, 1098, 1099, 1100, 1101, 1102, 1103, 1104, 1105, 1106, 1107, 1108, 1109, 1110, 1111, 1112, 1113, 1114, 1115, 1116, 1117, 1118, 1119, 1120, 1121, 1122, 1123, 1124, 1125, 1126, 1127, 1128, 1129, 1130, 1131, 1132, 1133, 1134, 1135, 1136, 1137, 1138, 1139, 1140, 1141, 1142, 1143, 1144, 1145, 1146, 1147, 1148, 1149, 1150, 1151, 1152, 1153, 1154, 1155, 1156, 1157, 1158, 1159, 1160, 1161, 1162, 1163, 1164, 1165, 1166, 1167, 1168, 1169, 1170, 1171, 1172, 1173, 1174, 1175, 1176, 1177, 1178, 1179, 1180, 1181, 1182, 1183, 1184, 1185, 1186, 1187, 1188, 1189, 1190, 1191, 1192, 1193, 1194, 1195, 1196, 1197, 1198, 1199, 1200, 1201, 1202, 1203, 1204, 1205, 1206, 1207, 1208, 1209, 1210, 1211, 1212, 1213, 1214, 1215, 1216, 1217, 1218, 1219, 1220, 1221, 1222, 1223, 1224, 1225, 1226, 1227, 1228, 1229, 1230, 1231, 1232, 1233, 1234, 1235, 1236, 1237, 1238, 1239, 1240, 1241, 1242, 1243, 1244, 1245, 1246, 1247, 1248, 1249, 1250, 1251, 1252, 1253, 1254, 1255, 1256, 1257, 1258, 1259, 1260, 1261, 1262, 1263, 1264, 1265, 1266, 1267, 1268, 1269, 1270, 1271, 1272, 1273, 1274, 1275, 1276, 1277, 1278, 1279, 1280, 1281, 1282, 1283, 1284, 1285, 1286, 1287, 1288, 1289, 1290, 1291, 1292, 1293, 1294, 1295, 1296, 1297, 1298, 1299, 1300, 1301, 1302, 1303, 1304, 1305, 1306, 1307, 1308, 1309, 1310, 1311, 1312, 1313, 1314, 1315, 1316, 1317, 1318, 1319, 1320, 1321, 1322, 1323, 1324, 1325, 1326, 1327, 1328, 1329, 1330, 1331, 1332, 1333, 1334, 1335, 1336, 1337, 1338, 1339, 1340, 1341, 1342, 1343, 1344, 1345, 1346, 1347, 1348, 1349, 1350, 1351, 1352, 1353, 1354, 1355, 1356, 1357, 1358, 1359, 1360, 1361, 1362, 1363, 1364, 1365, 1366, 1367, 1368, 1369, 1370, 1371, 1372, 1373, 1374, 1375, 1376, 1377, 1378, 1379, 1380, 1381, 1382, 1383, 1384, 1385, 1386, 1387, 1388, 1389, 1390, 1391, 1392, 1393, 1394, 1395, 1396, 1397, 1398, 1399, 1400, 1401, 1402, 1403, 1404, 1405, 1406, 1407, 1408, 1409, 1410, 1411, 1412, 1413, 1414, 1415, 1416, 1417, 1418, 1419, 1420, 1421, 1422, 1423, 1424, 1425, 1426, 1427, 1428, 1429, 1430, 1431, 1432, 1433, 1434, 1435, 1436, 1437, 1438, 1439, 1440, 1441, 1442, 1443, 1444, 1445, 1446, 1447, 1448, 1449, 1450, 1451, 1452, 1453, 1454, 1455, 1456, 1457, 1458, 1459, 1460, 1461, 1462, 1463, 1464, 1465, 1466, 1467, 1468, 1469, 1470, 1471, 1472, 1473, 1474, 1475, 1476, 1477, 1478, 1479, 1480, 1481, 1482, 1483, 1484, 1485, 1486, 1487, 1488, 1489, 1490, 1491, 1492, 1493, 1494, 1495, 1496, 1497, 1498, 1499, 1500, 1501, 1502, 1503, 1504, 1505, 1506, 1507, 1508, 1509, 1510, 1511, 1512, 1513, 1514, 1515, 1516, 1517, 1518, 1519, 1520, 1521, 1522, 1523, 1524, 1525, 1526, 1527, 1528, 1529, 1530, 1531, 1532, 1533, 1534, 1535, 1536, 1537, 1538, 1539, 1540, 1541, 1542, 1543, 1544, 1545, 1546, 1547, 1548, 1549, 1550, 1551, 1552, 1553, 1554, 1555, 1556, 1557, 1558, 1559, 1560, 1561, 1562, 1563, 1564, 1565, 1566, 1567, 1568, 1569, 1570, 1571, 1572, 1573, 1574, 1575, 1576, 1577, 1578, 1579, 1580, 1581, 1582, 1583, 1584, 1585, 1586, 1587, 1588, 1589, 1590, 1591, 1592, 1593, 1594, 1595, 1596, 1597, 1598, 1599, 1600, 1601, 1602, 1603, 1604, 1605, 1606, 1607, 1608, 1609, 1610, 1611, 1612, 1613, 1614, 1615, 1616, 1617, 1618, 1619, 1620, 1621, 1622, 1623, 1624, 1625, 1626, 1627, 1628, 1629, 1630, 1631, 1632, 1633, 1634, 1635, 1636, 1637, 1638, 1639, 1640, 1641, 1642, 1643, 1644, 1645, 1646, 1647, 1648, 1649, 1650, 1651, 1652, 1653, 1654, 1655, 1656, 1657, 1658, 1659, 1660, 1661, 1662, 1663, 1664, 1665, 1666, 1667, 1668, 1669, 1670, 1671, 1672, 1673, 1674, 1675, 1676, 1677, 1678, 1679, 1680, 1681, 1682, 1683, 1684, 1685, 1686, 1687, 1688, 1689, 1690, 1691, 1692, 1693, 1694, 1695, 1696, 1697, 1698, 1699, 1700, 1701, 1702, 1703, 1704, 1705, 1706, 1707, 1708, 1709, 1710, 1711, 1712, 1713, 1714, 1715, 1716, 1717, 1718, 1719, 1720, 1721, 1722, 1723, 1724, 1725, 1726, 1727, 1728, 1729, 1730, 1731, 1732, 1733, 1734, 1735, 1736, 1737, 1738, 1739, 1740, 1741, 1742, 1743, 1744, 1745, 1746, 1747, 1748, 1749, 1750, 1751, 1752, 1753, 1754, 1755, 1756, 1757, 1758, 1759, 1760, 1761, 1762, 1763, 1764, 1765, 1766, 1767, 1768, 1769, 1770, 1771, 1772, 1773, 1774, 1775, 1776, 1777, 1778, 1779, 1780, 1781, 1782, 1783, 1784, 1785, 1786, 1787, 1788, 1789, 1790, 1791, 1792, 1793, 1794, 1795, 1796, 1797, 1798, 1799, 1800, 1801, 1802, 1803, 1804, 1805, 1806, 1807, 1808, 1809, 1810, 1811, 1812, 1813, 1814, 1815, 1816, 1817, 1818, 1819, 1820, 1821, 1822, 1823, 1824, 1825, 1826, 1827, 1828, 1829, 1830, 1831, 1832, 1833, 1834, 1835, 1836, 1837, 1838, 1839, 1840, 1841, 1842, 1843, 1844, 1845, 1846, 1847, 1848, 1849, 1850, 1851, 1852, 1853, 1854, 1855, 1856, 1857, 1858, 1859, 1860, 1861, 1862, 1863, 1864, 1865, 1866, 1867, 1868, 1869, 1870, 1871, 1872, 1873, 1874, 1875, 1876, 1877, 1878, 1879, 1880, 1881, 1882, 1883, 1884, 1885, 1886, 1887, 1888, 1889, 1890, 1891, 1892, 1893, 1894, 1895, 1896, 1897, 1898, 1899, 1900, 1901, 1902, 1903, 1904, 1905, 1906, 1907, 1908, 1909, 1910, 1911, 1912, 1913, 1914, 1915, 1916, 1917, 1918, 1919, 1920, 1921, 1922, 1923, 1924, 1925, 1926, 1927, 1928, 1929, 1930, 1931, 1932, 1933, 1934, 1935, 1936, 1937, 1938, 1939, 1940, 1941, 1942, 1943, 1944, 1945, 1946, 1947, 1948, 1949, 1950, 1951, 1952, 1953, 1954, 1955, 1956, 1957, 1958, 1959, 1960, 1961, 1962, 1963, 1964, 1965, 1966, 1967, 1968, 1969, 1970, 1971, 1972, 1973, 1974, 1975, 1976, 1977, 1978, 1979, 1980, 1981, 1982, 1983, 1984, 1985, 1986, 1987, 1988, 1989, 1990, 1991, 1992, 1993, 1994, 1995, 1996,

100



Query Match 37.1%; Score 453.5; DB 2; Length 270;  
Best Local Similarity 42.5%; Pred. No. 5.8e-31;  
Matches 102; Conservative 36; Mismatches 87; Indels 15; Gaps 8;

Qy 2 LLSLTSLVLGSGWGGIPAIKPAKLSFSQRIYNGENAVLGSPWQVSLQ-DSSG---FH 58  
Db 1 MMLRLSSLLLVAVASGYGPPSSRP-----SSRVNGEDAVPYSWPQVSLQYKSGSPYH 56  
Qy 59 FCGSLISQSWVTAACHNCVSPGRHFVVLGEYDRSSNAEPLQVLSVRA--ITHPSWNST 116  
Db 57 TCGSLIADPWVVTAGHCISSTRYQVVLGEYDRVKEGPEQVPIPSNGDLFVPLWNRS 116  
Qy 117 TM--NNDTLLKLSAQYTRISPVCLASSNEALTEGLTCVTTGWGRLSGNGVNTPAHL 174  
Db 117 CVACGNDIALIKLSRAQLGDAVQLASLPAGDILPNETPIYITGWGLYINGPL-PPKL 175  
Qy 175 QOVALPLVTNNOCRO--YWDSSITDSMTICAGAGASSCGQSDGGLPLVC-QKGNVTWLIGI 231  
Db 176 QEALLPVVDYEHCSRWNWNGSVKRTMVCAGDIRSGCNGSDGGLPLNCPTEDGGQVHGV 235

RESULT 11  
A23473  
chymotrypsin-like proteinase (EC 3.4.21.-) - pig (tentative sequence) (fragments)  
N:Alternate names: pancreatic elastase II (misidentification)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 31-Mar-2000  
C:Accession: A23473  
R:Vered, M.; Gertler, A.; Burstein, Y.  
Int. J. Pept. Protein Res. 27, 183-190, 1986  
A:Reference number: A23473; MUID:86194934; PMID:3634756  
A:Accession: A23473  
A:Molecule type: protein  
A:Residues: 1-126 <VER>  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; serine proteinase

Query Match 36.9%; Score 451; DB 2; Length 126;  
Best Local Similarity 43.7%; Pred. No. 3.9e-31;  
Matches 93; Conservative 13; Mismatches 11; Indels 96; Gaps 3;

Qy 19 CGIPAIKPAKLSFSQRIYNGENAVLGSPWQVSLQDSSGFHFCGSLISQSWVTAACHNCV 78  
Db 1 CGVPAIKPALB-----IVBGZBAVPCSWPZWVSLZBSBGFHFCGSLISZ----- 45  
Qy 79 SPGRHFVVLGEYDRSSNAEPLQVLSVRAITHPSWNSTMMNDVTLKLSAQYTRIS 138  
Db 46 -----MNDLTLLKLSAQYTRIG 66  
Qy 139 PVCLASSNEALTEGLTCVTTGWGRLSGNGVNTPAHLQOVALPLVTNNOCROYWDSSTIDS 198  
Db 67 PVCLASGDEALPEXLPVCV-----MNDLTLLKLSAQYTRIG 84  
Qy 199 MICAGGASSCGQSDGGLPLVCQKGNVTWLIGI 231  
Db 85 MICAGASSCGQSDGGLPLVCERKGDXTWLXGI 117

RESULT 12  
B26823  
pancreatic elastase II (EC 3.4.21.71) A precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 16-Aug-1988 #sequence\_revision 16-Aug-1988 #text\_change 20-Jun-2000  
C:Accession: B26823; A27432; A41431; S34491  
R:Kawashima, I.; Tanai, T.; Shimoda, K.; Takiguchi, Y.  
DNA 6, 163-172, 1987  
A:Title: Characterization of pancreatic elastase II cDNAs: two elastase II mRNAs are expressed from a single gene  
A:Reference number: A90958; MUID:87217962; PMID:3646943  
A:Accession: B26823  
A:Molecule type: mRNA  
A:Residues: 1-269 <KAW>  
A:Cross-references: GB:M16652; NID:g182057; PIDN:AAA52380.1; PID:g182058

R:Fletcher, T.S.; Shen, W.F.; Largman, C.  
Biochemistry 26, 7256-7261, 1987  
A:Title: Primary structure of human pancreatic elastase 2 determined by sequence analysis  
A:Reference number: A27432; MUID:88107669; PMID:3427074  
A:Accession: A27432  
A:Molecule type: mRNA  
A:Residues: 1-269 <FLE>  
A:Cross-references: GB:M16631; NID:g182022; PIDN:AAA52374.1; PID:g182023  
R:Shirasu, Y.; Yoshida, H.; Matsuki, S.; Takemura, K.; Ikeda, N.; Shimada, Y.; Ozawa, J.  
Biochem. 102, 1555-1563, 1987  
A:Title: Molecular cloning and expression in *Escherichia coli* of a cDNA encoding human pancreatic elastase II  
A:Reference number: A41431; MUID:88198076; PMID:2834346  
A:Accession: A41431  
A:Molecule type: mRNA  
A:Residues: 1-201 'V', 203-269 <SHI>  
A:Cross-references: GB:D00236; NID:g219619; PIDN:BAA00165.1; PID:g219620  
A:Note: the authors translated the codon GTG for residue 202 as Cys  
R:Moulard, M.; Michon, T.; Kerfelec, B.; Chapus, C.  
FEBS Lett. 261, 179-183, 1990  
A:Title: Further studies on the human pancreatic binary complexes involving procarboxypeptidase Y  
A:Reference number: S08253; MUID:90169111; PMID:2307232  
A:Accession: S34491  
A:Molecule type: protein  
A:Residues: 'X', 18-50 <MOU>  
C:Genetics:  
A:Gene: GDB:ELAI  
A:Cross-references: GDB:I19866; OMIM:130120  
A:Map position: 12pter-12qter  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; pancreas; serine proteinase  
F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-28/Domain: propeptide #status predicted <PRO>  
F:29-269/Product: pancreatic elastase IIA #status predicted <NAT>  
F:29-262/Domain: trypsin homology <TRY>  
F:73-121,216/Active site: His, Asp, Ser #status predicted

Query Match 36.4%; Score 445; DB 2; Length 269;  
Best Local Similarity 43.9%; Pred. No. 3e-30;  
Matches 105; Conservative 28; Mismatches 92; Indels 14; Gaps 8;

Qy 2 LLSLTSLVLGSGWGGIPAIKPAKLSFSQRIYNGENAVLGSPWQVSLQDSSG---FH 58  
Db 1 MIRTLLSLTLAG-ALSCGDPTYP---YVTRVVGEEARPNSPWQVSLQYSSNGKWKYH 56  
Qy 59 FCGSLISQSWVTAACHNCVSPGRHFVVLGEYDRSSNAEPLQVLSVRAITHPSWNSTMM 118  
Db 57 TCGSLIANSWLTAACHLSSSTRYRVGLGRHNLVVAESGLAVSVKIVVHKDWNNSNOI 116  
Qy 119 N--NDVTLLKLSAQYTRISPVCLASSNEALTEGLTCVTTGWGRLSGNGVNTPAHLQO 176  
Db 117 SKGNDIALKLANPSLTDKIQACLPAGTILPNNYPCYVTVGWGLQTNGAV-PDV LQO 175  
Qy 177 VALPLVTNNOCRO---YWDSSITDSMICAGGAGA--SSCGSDGGLPLVCQKGN-TWVLIGI 231  
Db 176 GRLLVVDYATCGSSAWWGSSVKTSMTICAGDGVISSCNGSDGGLPLNCQASDGRQVHGI 234

RESULT 13  
ELRT2  
pancreatic elastase II (EC 3.4.21.71) precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 18-Aug-1982 #sequence\_revision 18-Aug-1982 #text\_change 24-Sep-1999  
C:Accession: A00961  
R:Macdonald, R.J.; Swift, G.H.; Quinto, C.; Swain, W.; Pictet, R.L.; Nikovits, W.; R  
Biochemistry 21, 1453-1463, 1982  
A:Title: Primary structure of two distinct rat pancreatic preproelastases determined by complementary DNA sequencing  
A:Reference number: A00960; MUID:82182967; PMID:6918221  
A:Accession: A00961  
A:Molecule type: mRNA  
A:Residues: 1-271 <MAC>  
A:Cross-references: GB:L00124; GB:J00731; NID:g204019; PIDN:AAA98780.1; PID:g204021  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; pancreas; serine proteinase; zymogen

F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-30/Domain: activation peptide #status predicted <APT>  
F:31-271/Product: elastase II #status predicted <MPT>  
F:31-264/Domain: trypsin homology <TRY>  
F:73,123,218/Active site: His, Asp, Ser #status predicted

Query Match 35.9%; Score 438; DB 1; Length 271;  
Best Local Similarity 41.8%; Pred. No. 1.2e-29;  
Matches 100; Conservative 36; Mismatches 91; Indels 12; Gaps 9;  
QY 2 LLLSLTSLVLLGSSWGGCIPAIKALPALSQRIVNGENAVLGSWPQVSLQD--SSG--PH 58  
DB 1 MIRTLLSLVAG--ALSCGPTYEYQHDVS--RVGQGEASPNWPQVSLQVLSGKWHH 58  
QY 59 FCGSLISQSWVTAACHCNVSPGRHVFVLGEYDRSSNAEPLOVLSVSRATIHPSWNSTTM 118  
DB 59 TCGGSLVANNWVLTAAHCLISRTYVLLGRHLSLTSSESGSLAVQVSKLVVHEKWNNAQL 118  
QY 119 N--NDVTLKLASPAQYTRISPVCLASNEALTEGLTCVTTGWGRLSGVGNVTPAHLOQ 176  
DB 119 SGNNDIALVKLASPVALTSKIOTACLPAGTILPNNYPCYVTGWGRLQTNQ--ATPDVLOQ 177  
QY 177 VALPLVTNQC--ROYWDSITDSMICAGGAG--ASSCGDSGGPLVCOKGN--TWVLIGI 231  
DB 178 GRLLVVDYATCSASWGSVTKTMVACAGDGVTSNCGDSGGPLNCQASNGQWVHGI 236

RESULT 14  
C26823  
pancreatic elastase II (EC 3.4.21.71) B precursor - human  
N:Alternate names: pancrreatopeptidase E  
C:Species: Homo sapiens (man)  
C:Date: 16-Aug-1988 #sequence\_revision 16-Aug-1988 #text\_change 22-Jun-1999  
C:Accession: C26823  
R:Kawashima, I.; Tani, T.; Shimoda, K.; Takiguchi, Y.  
DNA 6, 163-172, 1987  
A:Title: Characterization of pancreatic elastase II cDNAs: two elastase II mRNAs are expressed  
A:Reference number: A90958; MUID:87217962; PMID:3646943  
A:Accession: C26823  
A:Molecule type: mRNA  
A:Residues: 1-269 <RAW>  
A:Cross-references: GB:M16653; NID:g182059; PIDN:AAA52381.1; PID:g182060  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; serine proteinase  
F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-28/Domain: propeptide #status predicted <PRO>  
F:29-262/Product: pancreatic elastase IIB #status predicted <MAT>  
F:73,121,216/Active site: His, Asp, Ser #status predicted

Query Match 35.0%; Score 427; DB 2; Length 269;  
Best Local Similarity 41.0%; Pred. No. 1e-28;  
Matches 98; Conservative 36; Mismatches 91; Indels 14; Gaps 8;  
QY 2 LLLSLTSLVLLGSSWGGCIPAIKALPALSQRIVNGENAVLGSWPQVSLQDSSG---PH 58  
DB 1 MIRTLLSLVAG--ALSCGVTYAPDMS---RMLGGEARPNWPQVSLQVSSNGQWYH 56  
QY 59 FCGSLISQSWVTAACHCNVSPGRHVFVLGEYDRSSNAEPLOVLSVSRATIHPSWNSTTM 118  
DB 57 TCGGSLIANSWVLTAAHCLISRTYVLLGRHLSLTSSESGSLAVSIVKIVVHKDWSNQV 116  
QY 119 N--NDVTLKLASPAQYTRISPVCLASNEALTEGLTCVTTGWGRLSGVGNVTPAHLOQ 176  
DB 117 SKGNDIALKLANPVSLTDKIOLACLPAGTILPNNYPCYVTGWGRLQTNQ--ALPDDLQ 175  
QY 177 VALPLVTNQC--YWDSSITDSMICAGGAG--SSCGDSGGPLVCOKGN--TWVLIGI 231  
DB 176 GRLLVVDYATCSASWGSVTKTMNMCAGDGVICTNCGDSGGPLNCQASDGRWEVHGI 234

RESULT 15  
A29934

pancreatic elastase (EC 3.4.21.36) IIIA precursor - human  
N:Alternate names: protease E  
C:Species: Homo sapiens (man)  
C:Date: 15-Dec-1988 #sequence\_revision 15-Dec-1988 #text\_change 20-Jun-2000  
C:Accession: A29934; JX0045  
R:Tani, T.; Ohsumi, J.; Mita, K.; Takiguchi, Y.  
J. Biol. Chem. 263, 1231-1239, 1988  
A:Title: Identification of a novel class of elastase isozyme, human pancreatic elastase A  
A:Reference number: A92664; MUID:88087253; PMID:2826474  
A:Accession: A29934  
A:Molecule type: DNA  
A:Residues: 1-270 <TAN>  
A:Cross-references: GB:J03516  
R:Shirasu, Y.; Takemura, K.; Yoshida, H.; Sato, Y.; Iijima, H.; Shimada, Y.; Miyayama, J. Biochem. 104, 259-264, 1988  
A:Title: Molecular cloning of complementary DNA encoding one of the human pancreatic elastase A  
A:Reference number: JX0045; MUID:89034017; PMID:2460440  
A:Accession: JX0045

A:Molecule type: mRNA  
A:Residues: 1-240, G', 242-270 <SHI>  
A:Cross-references: GB:D00306; NID:g220013; PIDN:BAA00212.1; PID:g220014  
C:Comment: This enzyme is an alanine-specific serine proteinase that has little elast C:Genetics:  
C:Introns: 15/1; 43/3; 76/2; 121/2; 167/1; 214/3; 265/3  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; pancreas; serine proteinase  
F:1-15/Domain: signal sequence #status predicted <SIG>  
F:16-28/Domain: propeptide #status predicted <PRO>  
F:29-270/Product: pancreatic elastase IIIA #status experimental <MAT>  
F:29-263/Domain: trypsin homology <TRY>  
F:73,123,217/Active site: His, Asp, Ser #status predicted

Query Match 34.4%; Score 420.5; DB 2; Length 270;  
Best Local Similarity 40.7%; Pred. No. 3.6e-28;  
Matches 98; Conservative 35; Mismatches 91; Indels 17; Gaps 8;  
QY 1 MLLSLTSLVLLGSSWGGCIPAIKALPALSQRIVNGENAVLGSWPQVSLQ--DSSG--F 57  
DB 2 MLRLSSLLLVAVASGYG-----PPSHSSSRVHVHGEDAVPYSPWQVSLQTEKSGSFY 55  
QY 58 HFCGSLISQSWVTAACHCNVSPGRHVFVLGEYDRSSNAEPLOVLSV--RAITHPSWNS 115  
DB 56 HFCGSLIAPDWVVTAGHCISRDLYQVVLGEYNLAVKEGPEQVIFINSEELFVHPLWNR 115  
QY 116 TTM--NDVTLKLASPAQYTRISPVCLASNEALTEGLTCVTTGWGRLSGVGNVTPAH 173  
DB 116 SCVACNDIALKLSRAQLGDAVQLASLPAGDILPNKTPCYITCGRLYTNGLP--PDK 174  
QY 174 LQVALPLVTNQCRO--YWDSSITDSMICAGGAGSSCGDSGGPLVC--OKGNTWVLIG 230  
DB 175 LQARLPVVDYKHSRWNNWGSTVKTMTWCAGGYIRSGNCGDSGGPLNCTEDGGQVHVG 234  
QY 231 I 231  
DB 235 V 235

Search completed: December 20, 2002, 15:19:47  
Job time : 15.6087 secs





MOLECULE TYPE: No. 62324566  
US-08-944-483-62

Query Match 59.9%; Score 750.5; DB 4; Length 230;  
Best Local Similarity 58.0%; Pred. No. 4.3e-64;  
Matches 134; Conservative 35; Mismatches 61; Indels 1; Gaps 1;

QY 1 IVGENAVPGSWPQVSLQDNTGFHFCGGLISPNMVVTAACHQVTPGRHFVVLGEYDRS 60  
DB 1 IVGEDAVPGSWPQVSLQDNTGFHFCGGLISPNMVVTAACHQVTPGRHFVVLGEYDRS 59

QY 61 SNAEPQVLSIARAIHPNNANTMNDLTLKLASPARTAOVSPVCLASTNEALPSGL 120  
DB 60 SDENIQVLKIAKVPKPSILTVNNDITLLKLATPARFSQTVSAVCLPSADDDFPAGT 119

QY 121 TCVTGGRISGVNTVPARLQVPLVTVNOCROYWGARTIDAMICAGGSSCGQD 180  
DB 120 LCATTGCKTKYNANKTPDKLQQAALPLLSNAECKSKWGRRIIDVVICAGSGVSSCMGD 179

QY 181 SGGPLVCQKGNWTLIGIVSWGKNCNIQAPAMTRYKSFSTWVNOVMAYN 231  
DB 180 SGGPLVCQKGNWTLIGIVSWGSDTCTSSPGVYARVTKLIPWQKILAAAN 230

RESULT 2  
US-09-027-337-6  
Sequence 6, Application US/09027337B  
Patent No. 5972616  
GENERAL INFORMATION:  
APPLICANT: O'Brien, Timothy J.  
APPLICANT: Tanimoto, Hirotochi  
TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease Overexpressed in  
FILE REFERENCE: D6064  
CURRENT APPLICATION NUMBER: US/09/027, 337B  
PRIORITY FILING DATE: 1998-02-20  
NUMBER OF SEQ ID NOS: 13  
SEQ ID NO 6  
LENGTH: 231  
TYPE: PRT  
ORGANISM: Unknown  
FEATURE:  
OTHER INFORMATION: Serine protease catalytic domain of chymotrypsin (Chymb)  
OTHER INFORMATION: homologous to similar domain in TADG-15  
US-09-027-337-6

Query Match 59.9%; Score 750.5; DB 2; Length 231;  
Best Local Similarity 58.0%; Pred. No. 4.3e-64;  
Matches 134; Conservative 35; Mismatches 61; Indels 1; Gaps 1;

QY 1 IVGENAVPGSWPQVSLQDNTGFHFCGGLISPNMVVTAACHQVTPGRHFVVLGEYDRS 60  
DB 2 IVGEDAVPGSWPQVSLQDNTGFHFCGGLISPNMVVTAACHQVTPGRHFVVLGEYDRS 60

QY 61 SNAEPQVLSIARAIHPNNANTMNDLTLKLASPARTAOVSPVCLASTNEALPSGL 120  
DB 61 SDENIQVLKIAKVPKPSILTVNNDITLLKLATPARFSQTVSAVCLPSADDDFPAGT 120

QY 121 TCVTGGRISGVNTVPARLQVPLVTVNOCROYWGARTIDAMICAGGSSCGQD 180  
DB 120 LCATTGCKTKYNANKTPDKLQQAALPLLSNAECKSKWGRRIIDVVICAGSGVSSCMGD 180

QY 181 SGGPLVCQKGNWTLIGIVSWGKNCNIQAPAMTRYKSFSTWVNOVMAYN 231  
DB 181 SGGPLVCQKGNWTLIGIVSWGSDTCTSSPGVYARVTKLIPWQKILAAAN 231

RESULT 3  
US-09-644-600-6  
Sequence 6, Application US/09644600  
Patent No. 6451500  
GENERAL INFORMATION:  
APPLICANT: O'Brien, Timothy J.

APPLICANT: Tanimoto, Hirotochi  
TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease Overexpressed in Carcinomas  
FILE REFERENCE: D6064CIP/D  
CURRENT APPLICATION NUMBER: US/09/644, 600  
PRIORITY FILING DATE: 2000-08-23  
PRIORITY FILING DATE: 1999-10-20  
PRIORITY FILING DATE: 1998-02-20  
NUMBER OF SEQ ID NOS: 98  
SEQ ID NO 6  
LENGTH: 231  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Chymotrypsin  
US-09-644-600-6

Query Match 59.9%; Score 750.5; DB 4; Length 231;  
Best Local Similarity 58.0%; Pred. No. 4.3e-64;  
Matches 134; Conservative 35; Mismatches 61; Indels 1; Gaps 1;

QY 1 IVGENAVPGSWPQVSLQDNTGFHFCGGLISPNMVVTAACHQVTPGRHFVVLGEYDRS 60  
DB 2 IVGEDAVPGSWPQVSLQDNTGFHFCGGLISPNMVVTAACHQVTPGRHFVVLGEYDRS 60

QY 61 SNAEPQVLSIARAIHPNNANTMNDLTLKLASPARTAOVSPVCLASTNEALPSGL 120  
DB 61 SDENIQVLKIAKVPKPSILTVNNDITLLKLATPARFSQTVSAVCLPSADDDFPAGT 120

QY 121 TCVTGGRISGVNTVPARLQVPLVTVNOCROYWGARTIDAMICAGGSSCGQD 180  
DB 120 LCATTGCKTKYNANKTPDKLQQAALPLLSNAECKSKWGRRIIDVVICAGSGVSSCMGD 180

QY 181 SGGPLVCQKGNWTLIGIVSWGKNCNIQAPAMTRYKSFSTWVNOVMAYN 231  
DB 181 SGGPLVCQKGNWTLIGIVSWGSDTCTSSPGVYARVTKLIPWQKILAAAN 231

RESULT 4  
US-08-557-146-13  
Sequence 13, Application US/08557146  
Patent No. 5834290  
GENERAL INFORMATION:  
APPLICANT: Egelrud, Torbjorn  
APPLICANT: Hansson, Lennart  
TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic  
FILE REFERENCE: Enzyme (SCCE)  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: White & Case, Patent Department  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/557, 146  
FILING DATE: 14-DEC-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Sterner, Richard J.  
REGISTRATION NUMBER: 35,372  
REFERENCE/DOCKET NUMBER: 1103326-181  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 819-8783  
TELEFAX: (212) 354-8113



; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 229 amino acids  
; TYPE: amino acids  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: polypeptide  
US-08-557-146-13

Query Match 58.9%; Score 737; DB 2; Length 229;  
Best Local Similarity 57.6%; Pred. No. 8.3e-63;  
Matches 133; Conservative 36; Mismatches 60; Indels 2; Gaps 2;  
Qy 1 IVNGENAVPGSWPMQVSLQDNTGPHFCGSLISPNWVYTAACHCVTPGRHFVVLGEYDRS 60  
Db 1 IVNGEDAVPGSWPMQVSLQDNTGPHFCGSLISEDWVYTAACHGVRTS-DVVVAGEFDQG 59  
Qy 61 SNAEPVQVLSIARATHPHNWNANTMNDLTLKLASPARYTAQVSPVCLASTNEALPSGL 120  
Db 60 SDEENIQVLKIAKVKPKFSILTWNNDITLLKLPATPAREFQTVSAVCLPSADDDFPAGT 119  
Qy 121 TCVTTGWRISGVGNVTPARLQOVVPLVTVNQCRQYWGARTDAMICAGSGASSCGD 180  
Db 120 LCATTGWGKTKYNANKTPDKLQQAALPLLSNAECKKSGRRITDYMICA-GAGVSSCMGD 178  
Qy 181 SGGPLVCOKGNTWVLIGIVSMGCTKNCNIQAPAMYTRVSKFSTWVNOVMAYN 231  
Db 179 SGGPLVCOKGNTWVLIGIVSMGSDTCSTSSPGVYARVTKLIPWQKILAAAN 229

RESULT 5  
US-09-154-344-13  
; Sequence 13, Application US/09154344  
; Patent No. 5981256  
; GENERAL INFORMATION:  
; APPLICANT: Egelrud, Torbjorn  
; APPLICANT: Hansson, Lennart  
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic  
; TITLE OF INVENTION: Enzyme (SCCE)  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: White & Case, Patent Department  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2787  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/154,344  
; FILING DATE: 16-SEP-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/557,146  
; FILING DATE: 14-DEC-1995  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sterner, Richard J.  
; REGISTRATION NUMBER: 35,372  
; REFERENCE/DOCKET NUMBER: 1103326-181  
; TELEPHONE: (212) 819-8783  
; TELEFAX: (212) 354-8113  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 229 amino acids  
; TYPE: amino acids  
; STRANDEDNESS: single  
; TOPOLOGY: linear

; MOLECULE TYPE: polypeptide  
US-09-154-344-13

Query Match 58.9%; Score 737; DB 2; Length 229;  
Best Local Similarity 57.6%; Pred. No. 8.3e-63;  
Matches 133; Conservative 36; Mismatches 60; Indels 2; Gaps 2;  
Qy 1 IVNGENAVPGSWPMQVSLQDNTGPHFCGSLISPNWVYTAACHCVTPGRHFVVLGEYDRS 60  
Db 1 IVNGEDAVPGSWPMQVSLQDNTGPHFCGSLISEDWVYTAACHGVRTS-DVVVAGEFDQG 59  
Qy 61 SNAEPVQVLSIARATHPHNWNANTMNDLTLKLASPARYTAQVSPVCLASTNEALPSGL 120  
Db 60 SDEENIQVLKIAKVKPKFSILTWNNDITLLKLPATPAREFQTVSAVCLPSADDDFPAGT 119  
Qy 121 TCVTTGWRISGVGNVTPARLQOVVPLVTVNQCRQYWGARTDAMICAGSGASSCGD 180  
Db 120 LCATTGWGKTKYNANKTPDKLQQAALPLLSNAECKKSGRRITDYMICA-GAGVSSCMGD 178  
Qy 181 SGGPLVCOKGNTWVLIGIVSMGCTKNCNIQAPAMYTRVSKFSTWVNOVMAYN 231  
Db 179 SGGPLVCOKGNTWVLIGIVSMGSDTCSTSSPGVYARVTKLIPWQKILAAAN 229

RESULT 6  
US-08-278-091-10  
; Sequence 10, Application US/08278091  
; Patent No. 5506139  
; GENERAL INFORMATION:  
; APPLICANT: LOOSMORE, Sheena M  
; APPLICANT: YANG, Yan-Ping  
; APPLICANT: CHONG, Pele  
; APPLICANT: OOMEN, Raymond P.  
; APPLICANT: KLEIN, Michel H.  
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with  
; TITLE OF INVENTION: Reduced Protease Activity  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: Suite 701, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/278,091  
; FILING DATE: 21-JUL-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-371  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 228 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-278-091-10

Query Match 57.9%; Score 724.5; DB 1; Length 228;  
Best Local Similarity 57.1%; Pred. No. 1.3e-61;  
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;  
Qy 1 IVNGENAVPGSWPMQVSLQDNTGPHFCGSLISPNWVYTAACHCVTPGRHFVVLGEYDRS 60

Db 1 IVNGEAVPGSNPWQVSLQDKTGFEHFCGSLINENWVVTAAHCGVTTSDVYVAGEFDQ 59  
QY 61 SNAEPQVLSIARATHPNWNTNNDTLTKLSPARTYTAQVSPVCLASTNEALPSGL 120  
Db 60 SSSERIQKLIARVFNKNSKYNLSITNDITLTKLSTAAFSQTSVAVCLPSASDDFAAGT 119  
QY 121 TCVTGWRISGVNTPARLQOVVPLVTVNOCROYWGARTIDAMICAGSGASCCGD 180  
Db 120 TCVTGWRISGVNTPARLQOVVPLVTVNOCROYWGARTIDAMICAGSGASCCGD 177  
QY 181 SGGPLVCQKGNWTLGIVSGWTKNCNIQAPAMYTRVSKFSTWINGVMAYN 231  
Db 178 SGGPLVCCKNGAWTLGIVSGWSTCSTPGYARVATLVNWNQVOTLAAN 228

## RESULT 7

US-08-483-859-10  
; Sequence 10, Application US/08483859  
; Patent No. 5656436  
; GENERAL INFORMATION:  
; APPLICANT: LOOSMORE, Sheena M.  
; APPLICANT: YANG, Yan-Ping  
; APPLICANT: CHONG, Pele  
; APPLICANT: OOMEN, Raymond P.  
; APPLICANT: KLEIN, Michel H.  
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with  
; NUMBER OF INVENTIONS: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: Suite 701, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/483,859  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/296,149  
; FILING DATE: 26-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/278,091  
; FILING DATE: 21-JUL-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I.  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-493 MIS.vg  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 228 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-483-859-10

Query Match 57.9%; Score 724.5; DB 1; Length 228;  
Best Local Similarity 57.1%; Pred. No. 1.3e-61;  
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;  
QY 1 IVNGEAVPGSNPWQVSLQDKTGFEHFCGSLINENWVVTAAHCGVTTSDVYVAGEFDQ 59  
Db 1 IVNGEAVPGSNPWQVSLQDKTGFEHFCGSLINENWVVTAAHCGVTTSDVYVAGEFDQ 59

QY 61 SNAEPQVLSIARATHPNWNTNNDTLTKLSPARTYTAQVSPVCLASTNEALPSGL 120  
Db 60 SSSERIQKLIARVFNKNSKYNLSITNDITLTKLSTAAFSQTSVAVCLPSASDDFAAGT 119  
QY 121 TCVTGWRISGVNTPARLQOVVPLVTVNOCROYWGARTIDAMICAGSGASCCGD 180  
Db 120 TCVTGWRISGVNTPARLQOVVPLVTVNOCROYWGARTIDAMICAGSGASCCGD 177  
QY 181 SGGPLVCQKGNWTLGIVSGWTKNCNIQAPAMYTRVSKFSTWINGVMAYN 231  
Db 178 SGGPLVCCKNGAWTLGIVSGWSTCSTPGYARVATLVNWNQVOTLAAN 228

## RESULT 8

US-08-472-173-10  
; Sequence 10, Application US/08472173  
; Patent No. 5665353  
; GENERAL INFORMATION:  
; APPLICANT: LOOSMORE, Sheena M.  
; APPLICANT: YANG, Yan-Ping  
; APPLICANT: CHONG, Pele  
; APPLICANT: OOMEN, Raymond P.  
; APPLICANT: KLEIN, Michel H.  
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with  
; NUMBER OF INVENTIONS: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: Suite 701, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/472,173  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/296,149  
; FILING DATE: 26-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/278,091  
; FILING DATE: 21-JUL-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I.  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-493 MIS.vg  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 228 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-472-173-10

Query Match 57.9%; Score 724.5; DB 1; Length 228;  
Best Local Similarity 57.1%; Pred. No. 1.3e-61;  
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;  
QY 1 IVNGEAVPGSNPWQVSLQDKTGFEHFCGSLINENWVVTAAHCGVTTSDVYVAGEFDQ 59  
Db 1 IVNGEAVPGSNPWQVSLQDKTGFEHFCGSLINENWVVTAAHCGVTTSDVYVAGEFDQ 59  
QY 61 SNAEPQVLSIARATHPNWNTNNDTLTKLSPARTYTAQVSPVCLASTNEALPSGL 120

Db 60 SSSEKIQLKIAKFKNSKYNLSLTINNDITLLKLSAAESFQTSVAVCLPSASDDFAAGT 119  
Qy 121 TCVTTGWRISGVGNVTTPARLQOVVPLVTVNQCROYWGARITDAMICAGSGASSCGD 180  
Db 120 TCVTTGWR-LTRYAN-TPDRIQQAASLPLLSNTNCKKYGWGTIKIDAMICAGASGVSSCMGD 177  
Qy 181 SGGPLVCOKGNTWVLIGVSGTKNCNQAPAMYTRVSKFSTWVNOVMAYN 231  
Db 178 SGGPLVCKKNGAWTLVIGVSGSSCTSTPGVYARVTVLNVWVQOTLAAN 228

RESULT 9  
US-08-487-167-10  
: Sequence 10, Application US/08487167  
: Patent No. 5869302  
: GENERAL INFORMATION:  
: APPLICANT: LOOSMORE, Sheena M.  
: APPLICANT: YANG, Yan-Ping  
: APPLICANT: CHONG, Pele  
: APPLICANT: OOMEN, Raymond P.  
: APPLICANT: KLEIN, Michel H.  
: TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with  
: TITLE OF INVENTION: Reduced Protease Activity  
: NUMBER OF SEQUENCES: 23  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Sim & McBurney  
: STREET: Suite 701, 330 University Avenue  
: CITY: Toronto  
: STATE: Ontario  
: COUNTRY: Canada  
: ZIP: M5G 1R7  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/487.167  
: FILING DATE: 07-JUN-1995  
: CLASSIFICATION: 435  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/296,149  
: FILING DATE: 26-AUG-1994  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/278,091  
: FILING DATE: 21-JUL-1994  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Stewart, Michael I.  
: REGISTRATION NUMBER: 24,973  
: REFERENCE/DOCKET NUMBER: 1038-508 MIS:vg  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (416) 595-1155  
: INFORMATION FOR SEQ ID NO: 10:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 228 amino acids  
: TYPE: amino acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
US-08-487-167-10

Query Match 57.9%; Score 724.5; DB 2; Length 228;  
Best Local Similarity 57.1%; Pred. No. 1.3e-61;  
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;  
Qy 1 IVNGENAVPGSWPQVSLQDNTGFHFCGSLSPNNVYTAACQVTPGRHFFVLGEYDRS 60  
Db 1 IVNGEAVPGSWPQVSLQDNTGFHFCGSLSPNNVYTAACQVTPGRHFFVLGEYDRS 59  
Qy 61 SNAEPVQLSTARATHHPNNWNTNNDITLLKLSAPRYTAQVSPVCLASTNEALPSGL 120  
Db 60 SSSEKIQLKIAKFKNSKYNLSLTINNDITLLKLSAAESFQTSVAVCLPSASDDFAAGT 119

Qy 121 TCVTTGWRISGVGNVTTPARLQOVVPLVTVNQCROYWGARITDAMICAGSGASSCGD 180  
Db 120 TCVTTGWR-LTRYAN-TPDRIQQAASLPLLSNTNCKKYGWGTIKIDAMICAGASGVSSCMGD 177  
Qy 181 SGGPLVCOKGNTWVLIGVSGTKNCNQAPAMYTRVSKFSTWVNOVMAYN 231  
Db 178 SGGPLVCKKNGAWTLVIGVSGSSCTSTPGVYARVTVLNVWVQOTLAAN 228

RESULT 10  
US-08-482-816-10  
: Sequence 10, Application US/08482816  
: Patent No. 5935573  
: GENERAL INFORMATION:  
: APPLICANT: LOOSMORE, Sheena M  
: APPLICANT: YANG, Yan-Ping  
: APPLICANT: CHONG, Pele  
: APPLICANT: OOMEN, Raymond P.  
: APPLICANT: KLEIN, Michel H.  
: TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with Reduced Protease A  
: NUMBER OF SEQUENCES: 23  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Sim & McBurney  
: STREET: Suite 701, 330 University Avenue  
: CITY: Toronto  
: STATE: Ontario  
: COUNTRY: Canada  
: ZIP: M5G 1R7  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/482.816  
: FILING DATE: 07-JUN-1995  
: CLASSIFICATION: 435  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/296,149  
: FILING DATE: 26-AUG-1994  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/278,091  
: FILING DATE: 21-JUL-1994  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Stewart, Michael I.  
: REGISTRATION NUMBER: 24,973  
: REFERENCE/DOCKET NUMBER: 1038-494 MIS:vg  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (416) 595-1155  
: TELEFAX: (416) 595-1163  
: INFORMATION FOR SEQ ID NO: 10:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 228 amino acids  
: TYPE: amino acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
US-08-482-816-10

Query Match 57.9%; Score 724.5; DB 2; Length 228;  
Best Local Similarity 57.1%; Pred. No. 1.3e-61;  
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;  
Qy 1 IVNGENAVPGSWPQVSLQDNTGFHFCGSLSPNNVYTAACQVTPGRHFFVLGEYDRS 60  
Db 1 IVNGEAVPGSWPQVSLQDNTGFHFCGSLSPNNVYTAACQVTPGRHFFVLGEYDRS 59  
Qy 61 SNAEPVQLSTARATHHPNNWNTNNDITLLKLSAPRYTAQVSPVCLASTNEALPSGL 120  
Db 60 SSSEKIQLKIAKFKNSKYNLSLTINNDITLLKLSAAESFQTSVAVCLPSASDDFAAGT 119  
Qy 121 TCVTTGWRISGVGNVTTPARLQOVVPLVTVNQCROYWGARITDAMICAGSGASSCGD 180

Db 120 TCVTGNG-LTRYAN-TPDLQASLPLLSNTNCKYWGTKIKDAMICAGASGVSSCMGD 177  
QY 181 SGGPLVCQKGNWTWLVIGISWGTKNQCIQAPAMYTRVSKFSTWVNOVMAYN 231  
Db 178 SGGPLVCCKNGAWTLVIGISWGSTCSTPGVYARVATLVNWWQOTLAAN 228

## RESULT 11

US-08-296-149-10 Application US/08296149  
; Sequence 10, Application US/08296149  
; Patent No. 593297  
; GENERAL INFORMATION:  
; APPLICANT: LOOSMORE, Sheena M  
; APPLICANT: YANG, Yan-Ping  
; APPLICANT: CHONG, Pele  
; APPLICANT: OOMEN, Raymond P.  
; APPLICANT: KLEIN, Michel H.  
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: Suite 701, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/296,149  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-390  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 228 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-296-149-10

Query Match 57.9%; Score 724.5; DB 2; Length 228;  
Best Local Similarity 57.1%; Pred. No. 1.3e-61;  
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;

QY 1 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVYTAHCQVTPGRHFVVLGEYDRS 60  
Db 1 IVNGEEAVPGSWPQVSLQDNTGFHFCGSLINENWVYTAHCGVTTs-DVVVAGEFDQ 59  
QY 61 SNAEPQVLSIARLTHPNWNTMNDLTLKIASPARVTAQVSPVCLASTNEALPSGL 120  
Db 60 SSSEKIQKLIARVFNKNSYNSLTINNDITLLKLSAASFQSVAVCLPSASDDFAAGT 119  
QY 121 TCVTGNGRISGVGNTVPARLQOVLPVTVNOCROYWGARTIDAMICAGSGASSCOGD 180  
Db 120 TCVTGNG-LTRYAN-TPDLQASLPLLSNTNCKYWGTKIKDAMICAGASGVSSCMGD 177  
QY 181 SGGPLVCQKGNWTWLVIGISWGTKNQCIQAPAMYTRVSKFSTWVNOVMAYN 231  
Db 178 SGGPLVCCKNGAWTLVIGISWGSTCSTPGVYARVATLVNWWQOTLAAN 228

## RESULT 12

US-08-801-499-10  
; Sequence 10, Application US/08801499  
; Patent No. 5962430  
; GENERAL INFORMATION:  
; APPLICANT: LOOSMORE, Sheena M  
; APPLICANT: YANG, Yan-Ping  
; APPLICANT: CHONG, Pele  
; APPLICANT: OOMEN, Raymond P.  
; APPLICANT: KLEIN, Michel H.  
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with Reduced Protease A  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: 6th Floor, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/801,499  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/482,816  
; FILING DATE: 07-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/296,149  
; FILING DATE: 26-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/278,091  
; FILING DATE: 21-JUL-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-671 MIS:jb  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 228 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-801-499-10

Query Match 57.9%; Score 724.5; DB 2; Length 228;  
Best Local Similarity 57.1%; Pred. No. 1.3e-61;  
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;

QY 1 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVYTAHCQVTPGRHFVVLGEYDRS 60  
Db 1 IVNGEEAVPGSWPQVSLQDNTGFHFCGSLINENWVYTAHCGVTTs-DVVVAGEFDQ 59  
QY 61 SNAEPQVLSIARLTHPNWNTMNDLTLKIASPARVTAQVSPVCLASTNEALPSGL 120  
Db 60 SSSEKIQKLIARVFNKNSYNSLTINNDITLLKLSAASFQSVAVCLPSASDDFAAGT 119  
QY 121 TCVTGNGRISGVGNTVPARLQOVLPVTVNOCROYWGARTIDAMICAGSGASSCOGD 180  
Db 120 TCVTGNG-LTRYAN-TPDLQASLPLLSNTNCKYWGTKIKDAMICAGASGVSSCMGD 177  
QY 181 SGGPLVCQKGNWTWLVIGISWGTKNQCIQAPAMYTRVSKFSTWVNOVMAYN 231  
Db 178 SGGPLVCCKNGAWTLVIGISWGSTCSTPGVYARVATLVNWWQOTLAAN 228

## RESULT 13

```

US-08-615-271-10
; Sequence 10, Application US/08615271
; Patent No. 5981503
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED
; TITLE OF INVENTION: PROTEASE ACTIVITY
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/615,271
; FILING DATE: 20-JUN-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-580
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-615-271-10

```

```

Query Match 57.9%; Score 724.5; DB 2; Length 228;
Best Local Similarity 57.1%; Pred. No. 1.3e-61;
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;

Qy 1 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVYTAARHCQVTPGRHFVVLGEYDRS 60
Db 1 IVNGEAVPGSWPQVSLQDNTGFHFCGSLINENWVYTAARHCQVTS-DVYVAGEFDQG 59

Qy 61 SNAEPVQVLSARATHPHNNANTMNDLTLKLSAPARYTAQVSPVCLASTNEALPSGL 120
Db 60 SSSEKIQKLAKVFKNSKYNLSLTINDITLLKLSAASFSQTSVAVCLPSASDDFAAGT 119

Qy 121 TCVTTGWRISGVGNVTPARLQVVLPLVTNQCQRYNGARITDAMICAGSGASCCGD 180
Db 120 TCVTTGWR-LTRYAN-TDRLQOASLPLLSNTNCKYNGTKIKDAMICAGSGVSSCMGD 177

Qy 181 SGGPLVCKKNGAWTLVIGVSGWTKNCNTQAPAMYTRVSKFSTWVQVMAYN 231
Db 178 SGGPLVCKKNGAWTLVIGVSGWSSSTCSTPGYARVYALVNVVQOTLAAN 228

RESULT 14
US-09-074-660-10
; Sequence 10, Application US/09074660
; Patent No. 6020183
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.

```

```

; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; TITLE OF INVENTION: Reduced Protease Activity
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/074,660
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,167
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-731 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-074-660-10

Query Match 57.9%; Score 724.5; DB 3; Length 228;
Best Local Similarity 57.1%; Pred. No. 1.3e-61;
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;

Qy 1 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVYTAARHCQVTPGRHFVVLGEYDRS 60
Db 1 IVNGEAVPGSWPQVSLQDNTGFHFCGSLINENWVYTAARHCQVTS-DVYVAGEFDQG 59

Qy 61 SNAEPVQVLSARATHPHNNANTMNDLTLKLSAPARYTAQVSPVCLASTNEALPSGL 120
Db 60 SSSEKIQKLAKVFKNSKYNLSLTINDITLLKLSAASFSQTSVAVCLPSASDDFAAGT 119

Qy 121 TCVTTGWRISGVGNVTPARLQVVLPLVTNQCQRYNGARITDAMICAGSGASCCGD 180
Db 120 TCVTTGWR-LTRYAN-TDRLQOASLPLLSNTNCKYNGTKIKDAMICAGSGVSSCMGD 177

Qy 181 SGGPLVCKKNGAWTLVIGVSGWTKNCNTQAPAMYTRVSKFSTWVQVMAYN 231
Db 178 SGGPLVCKKNGAWTLVIGVSGWSSSTCSTPGYARVYALVNVVQOTLAAN 228

RESULT 15
US-09-074-659-10
; Sequence 10, Application US/09074659
; Patent No. 6025342
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele

```

APPLICANT: OOMEN, Raymond P.  
APPLICANT: KLEIN, Michel H.  
TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with  
TITLE OF INVENTION: Reduced Protease Activity  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 6th Floor, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/074,659  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/487,167  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/296,149  
FILING DATE: 26-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/278,091  
FILING DATE: 21-JUL-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I.  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-730 MIS:jb  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 228 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-074-659-10

Query Match 57.9%; Score 724.5; DB 3; Length 228;  
Best Local Similarity 57.1%; Pred. No. 1.3e-61;  
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;  
QY 1 IVNGENAVPGSWPQVSLQDNTGFHFCGGLSPNNVVTAAHCQVTPGRHFVVLGEYDRS 60  
Db 1 IVNGEAVPGSWPQVSLQDNTGFHFCGGLSPNNVVTAAHCQVTPGRHFVVLGEYDRS 60  
QY 61 SNAEPQVLSIARATHPNMNTMNDITLLKLASPARYTAQVSPVCLASTNEALPSGL 120  
Db 60 SSSEKIQKLIARVFNKYNLSLTINNDITLLKLSTAAFSQTSVAVCLPSASDDFAAGT 119  
QY 121 TCVTTGWRISGVNTPARLQOVLPVTVNOCROGVARITDAMICAGSGSCQGD 180  
Db 120 TCVTTGWRISGVNTPARLQOVLPVTVNOCROGVARITDAMICAGSGSCQGD 177  
QY 181 SGGPLVCKNGAWTLVIGVSWGSKNIOAPMYTRVSKFSTWVNOVMAYN 231  
Db 178 SGGPLVCKNGAWTLVIGVSWGSKNIOAPMYTRVSKFSTWVNOVMAYN 228

Search completed: December 20, 2002, 15:20:39  
Job time : 14.087 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:12:58 ; Search time 7.91304 Seconds  
(without alignments)  
1210.790 Million cell updates/sec

Title: US-09-856-319B-4\_COPY\_34\_264

Perfect score: 1252

Sequence: 1 IVNGENAVPGSWPQVSLQD.....AMYTRVSKFSTWVQVWVAYN 231

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1119	89.4	264	1	CTRL_HUMAN
2	757.5	60.5	263	1	CTRL_CANFA
3	750.5	59.9	263	1	CTRL_GADMO
4	750.5	59.9	263	1	CTRL_HUMAN
5	740.5	59.1	245	1	CTRL_BOVIN
6	727.5	58.1	245	1	CTRL_BOVIN
7	715.5	57.1	263	1	CTRL_RAT
8	714.5	57.1	245	1	CTRL_GADMO
9	503	40.2	269	1	EL2_PIG
10	499	39.9	271	1	EL2_RAT
11	497.5	39.7	253	1	CAC3_BOVIN
12	486	38.8	269	1	EL2A_HUMAN
13	481.5	38.5	270	1	EL3B_HUMAN
14	480	38.3	269	1	EL2_BOVIN
15	475	37.9	271	1	EL2_MOUSE
16	464.5	37.1	810	1	PLMN_HUMAN
17	462	36.9	812	1	PLMN_MOUSE
18	459.5	36.7	270	1	EL3A_HUMAN
19	455.5	36.4	270	1	TRYT_MERUN
20	455.5	36.4	342	1	PS8_RAT
21	455	36.3	454	1	TMS3_HUMAN
22	452.5	36.1	275	1	TRYT_PIG
23	452.5	36.1	343	1	PLMN_SHEEP
24	451.5	36.1	810	1	PLMN_MACMU
25	449.5	35.9	790	1	PLMN_PIG
26	449	35.9	269	1	EL2B_HUMAN
27	448.5	35.8	333	1	PLMN_HUMAN
28	448.5	35.8	343	1	PS8_HUMAN
29	447.5	35.7	268	1	CLCR_HUMAN
30	447.5	35.7	268	1	CLCR_RAT
31	447.5	35.7	273	1	MCT7_MOUSE
32	447.5	35.7	342	1	PS8_MOUSE
33	444	35.5	231	1	TRYP_PIG

34	442.5	35.3	273	1	MCT7_RAT	P27435 rattus norv
35	440	35.1	246	1	TRY1_RAT	P00762 rattus norv
36	439.5	35.1	274	1	MCT6_RAT	P50343 rattus norv
37	439.5	35.1	490	1	TMS2_MOUSE	Q9J1q8 mus musculu
38	439	35.1	248	1	TRY1_CHICK	P80010 gallus gall
39	437.5	34.9	338	1	PLMN_HORSE	P06871 canis fami
40	436	34.8	246	1	TRY1_CANFA	P06871 canis fami
41	435	34.7	248	1	TRY2_CHICK	Q90628 gallus gall
42	433	34.6	247	1	TRY2_CANFA	P08872 canis fami
43	431	34.4	246	1	TRY2_RAT	P00773 rattus norv
44	430.5	34.4	266	1	EL1_RAT	P21845 mus musculu
45	430.5	34.4	276	1	MCT6_MOUSE	

#### ALIGNMENTS

RESULT 1	CTRL_HUMAN	STANDARD;	PRT;	264 AA.
ID	CTRL_HUMAN			
AC	P40313;			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Chymotrypsin-like protease CTRL-1 precursor (EC 3.4.21.-).			
GN	CTRL OR CTRL1			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=94093544; PubMed=8268911;			
RA	Larsen F., Solheim J., Kristensen T., Kolsto A.B., Prydz H.;			
RT	"A tight cluster of five unrelated human genes on chromosome 16q22.1."			
RL	Hum. Mol. Genet. 2:1589-1595(1993).			
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X71874; CAA50710.1; -			
DR	EMBL; X71877; CAA50711.1; -			
DR	HSSP; P00763; IDPO.			
DR	MEROPS; S01.256; -			
DR	Genew; HGNC:2524; CTRL.			
DR	MIM; 118888; -			
DR	InterPro; IPR001314; Chymotrypsin.			
DR	InterPro; IPR001254; Ser_protease_Try.			
DR	Pfam; PF00089; trypsin; 1.			
DR	PRINTS; P00722; CHYMOTRYPSIN.			
DR	SMART; SM00020; TRYP_SPC; 1.			
DR	PROSITE; PS0240; TRYP_SIN_DOM; 1.			
DR	PROSITE; PS00134; TRYP_SIN_HIS; 1.			
DR	PROSITE; PS00135; TRYP_SIN_SER; 1.			
KW	Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal.			
FT	SIGNAL 1 18 POTENTIAL.			
FT	PROPEP 19 33*			
FT	CHAIN 34 264			
FT	ACT_SITE 75 75			
FT	ACT_SITE 121 121			
FT	ACT_SITE 214 214			
FT	CARBOHYD 114 114			
FT	DISULFID 19 141			
FT	DISULFID 60 76			
FT	DISULFID 155 220			
FT	DISULFID 187 201			

FT	CHAIN	34	164	CHYMOTRYPSIN 2, B CHAIN.
FT	CHAIN	167	263	CHYMOTRYPSIN 2, C CHAIN.
FT	ACT_SITE	75	75	CHARGE RELAY SYSTEM.
FT	ACT_SITE	120	120	CHARGE RELAY SYSTEM.
FT	ACT_SITE	213	213	CHARGE RELAY SYSTEM.
FT	DISULFID	19	140	BY SIMILARITY.
FT	DISULFID	60	76	BY SIMILARITY.
FT	DISULFID	154	219	BY SIMILARITY.
FT	DISULFID	186	200	BY SIMILARITY.
FT	DISULFID	209	238	BY SIMILARITY.
SQ	SEQUENCE	263 AA;	27787 MR;	2A2F449D813B3961 CRC64;

Query Match 60.5%; Score 757.5; DB 1; Length 263;  
 Best Local Similarity 58.0%; Pred. No. 7.7e-61;  
 Matches 134; Conservative 35; Mismatches 61; Indels 1; Gaps

QY	1	IVNGENAVPGSWPQVWSLQDNTGFHFCGGSLSIPNNVVTAAHCQVTPGRHFVVLGEYDKS	60
DB	34	IVNGEDAVPGSWPQVWSLQDNTGFHFCGGSLSIEDWVVTAAHCGVRT-THQVYVAGEFDQ	99
QY	61	SNAPVQVLSIARAIATHPNNNANTMNDLTLLKLSPARYTAQVSPVCLASTNEALPSGL	120
DB	93	SDAIEQVLKIAKVFKNPKFMFTINDDITLLKLATPAFESKTVSAVCLPQATDDFPAGT	152
QY	121	TCVTTGWRISGVGNVTPARLOQVVLPLVTVNCQYWCAGRIITAMICAGSGASCQGD	180
DB	153	LCVTTGWLTKHTNANTPDKLQQAALLPSNAECKKFWGSKITDLMVCGAGSGVSCMGD	212
QY	181	SGGPLVCQKQNTWVLGIVSWGKTKNCNIQAPAMYTRVSKFSTWINQVMAYN	231
DB	213	SGGPLVCQKQDAWTLGIVSWGTCSTSTPGVYARVTKLIPVWQIILQAN	263

RESULT 3

ID	CTRA_GADMO	STANDARD;	PRT;	263 AA.
AC	P47796;			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Chymotrypsin A precursor (EC 3.4.21.1).			
OS	Gadus morhua Atlantic cod.			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;			
OC	Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.			
OX	NCBI_Taxid=8049;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Pyloric caeca;			
RX	MEDLINE=94368860; PubMed=8086467;			
RA	Gudmundsdottir A., Oskarsson S., Eakin A.E., Craik C.S.,			
RA	Bjarnason J.B.;			
RT	"Atlantic cod cDNA encoding a psychrophilic chymotrypsinogen.";			
RL	Biochim. Biophys. Acta 1219:211-214(1994).			
RN	[2]			
RP	SEQUENCE OF 19-30 AND 34-49.			
RC	TISSUE=Pyloric caeca;			
RX	MEDLINE=92111252; PubMed=1764912;			
RA	Asgerlsson B., Bjarnason J.B.;			
RT	"Structural and kinetic properties of chymotrypsin from Atlantic cod			
RL	(Gadus morhua). Comparison with bovine chymotrypsin.";			
RL	Comp. Biochem. Physiol. 99B:327-335(1991).			
CC	-1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,			
CC	Phe-I-Xaa, Leu-I-Xaa.			
CC	-1- SUBCELLULAR LOCATION: Extracellular.			
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.			

This SWISS-PROT entry is copyright. It is produced through a collabor between the Swiss Institute of Bioinformatics and the EMBL outstat the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no modified and this statement is not removed. Usage by and for comm entities requires a license agreement (See <http://www.isb-sib.ch/anno>



or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch).

```
CC -----
CC EMBL; X78490; CAAS5242.1; -.
CC DR HSP; P00766; ICHG.
CC DR MEROPS; S01.152; -.
CC DR InterPro; IPR001314; Chymotrypsin.
CC DR InterPro; IPR001254; Ser_protease_Try.
CC DR Pfam; PF00089; trypsin; 1.
CC DR PRINTS; PR00722; CHYMOTRYPSIN.
CC DR SMART; SM00020; Tryp_SPC; 1.
CC DR PROSITE; PS00240; TRYPSIN_DOM; 1.
CC DR PROSITE; PS00134; TRYPSIN_HIS; 1.
CC DR PROSITE; PS00135; TRYPSIN_SER; 1.
CC KW Hydrolyase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
CC FT SIGNAL 1 18
CC FT CHAIN 19 263 CHYMOTRYPSIN A.
CC FT ACT_SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT ACT_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT ACT_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT DISULFID 19 140 BY SIMILARITY.
CC FT DISULFID 60 76 BY SIMILARITY.
CC FT DISULFID 154 219 BY SIMILARITY.
CC FT DISULFID 186 200 BY SIMILARITY.
CC FT DISULFID 209 238 BY SIMILARITY.
CC FT CONFLICT 21 21 R -> S (IN REF. 2).
CC FT CONFLICT 25 25 S -> O (IN REF. 2).
CC FT CONFLICT 29 29 T -> S (IN REF. 2).
CC FT CONFLICT 44 44 S -> T (IN REF. 2).
CC FT CONFLICT 46 46 S -> Y (IN REF. 2).
CC SQ SEQUENCE 263 AA; 28294 MW; 47AAC699A0A64FBB CRC64;

Query Match 59.9%; Score 750.5; DB 1; Length 263;
Best Local Similarity 58.4%; Pred. No. 3.3e-60;
Matches 135; Conservative 33; Mismatches 62; Indels 1; Gaps 1;

QY 1 IVNGENAVPGSWPQVSLQDNTGFHCGGSLISPNWVYTAACHQVTPGRHFVVLGEYDYS 60
DB 34 IVNGEAVPHSWQVSLQDQTFHFCGGLINENWVYTAACHNV-KNYHRVVLGEHDS 92
QY 61 SNAEPQVLTARAIHPNNANTMNDLTLKLLASPARYTAQVSPVCLASTNEALPSGL 120
DB 93 SNSEGVQVMTVGQVFKRHPNGFTINDILLVKLATPATLNMVSPVCLAEITDDVFE 152
QY 121 TCVTGTGGRISGVGNVTPARLQVVLPLVTNOCROYGARITDAMICAGSGASSCGD 180
DB 153 KVTSGWGLTRYNAADTPALLOQAALPLLTNEQCKKFWGNKISDLMICAGAGASS 212
QY 181 SGGPLVCQKAGNTWVLIGIVSGMGTKNICQAPAMVTRYSKFSTWYNQVAYN 231
DB 213 SGGPLVCQKAGSWTLVIGVSGSGCTCTPTMPGVYARVTELRAWVDQTAAAN 263
```

## RESULT 4

```
CTRB_HUMAN STANDARD; PRT; 263 AA.
ID CTRB_HUMAN
AC P17538;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymotrypsinogen B precursor (EC 3.4.21.1).
GN CTRB1 OR CTRB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=89134264; PubMed=2917002;
RA Tomita N., Izumoto Y., Horii A., Doi S., Yokouchi H., Ogawa M.,
RA Mori T., Matsubara K.;
RT "Molecular cloning and nucleotide sequence of human pancreatic
prechymotrypsinogen cDNA.";
```

```
RL Biochem. Biophys. Res. Commun. 158:569-575(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,
CC Phe-I-Xaa, Leu-I-Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; M24400; AAA52128.1; -.
CC EMBL; BC005385; AAH05385.1; -.
CC PIR; A31299; A31299.
CC HSP; P00766; ICHG.
CC MEROPS; S01.152; -.
CC Genew; HGNC:2521; CTRB1.
CC MIM; 118890; -.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC SMART; SM00020; Tryp_SPC; 1.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC KW Hydrolyase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
CC FT SIGNAL 1 18
CC FT CHAIN 19 263 CHYMOTRYPSIN B.
CC FT CHAIN 19 31 CHYMOTRYPSIN B, A CHAIN.
CC FT CHAIN 34 164 CHYMOTRYPSIN B, B CHAIN.
CC FT CHAIN 167 263 CHYMOTRYPSIN B, C CHAIN.
CC FT ACT_SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT ACT_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT ACT_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT DISULFID 19 140 BY SIMILARITY.
CC FT DISULFID 60 76 BY SIMILARITY.
CC FT DISULFID 154 219 BY SIMILARITY.
CC FT DISULFID 186 200 BY SIMILARITY.
CC FT DISULFID 209 238 BY SIMILARITY.
CC SQ SEQUENCE 263 AA; 27870 MW; 4C1C055A490B8701 CRC64;

Query Match 59.9%; Score 750.5; DB 1; Length 263;
Best Local Similarity 58.0%; Pred. No. 3.3e-60;
Matches 134; Conservative 35; Mismatches 61; Indels 1; Gaps 1;

QY 1 IVNGENAVPGSWPQVSLQDNTGFHCGGSLISPNWVYTAACHQVTPGRHFVVLGEYDYS 60
DB 34 IVNGEDAVPGSWPQVSLQDQTFHFCGGLISEDWVYTAACHGVRTS-DVVAGEFDQG 92
QY 61 SNAEPQVLTARAIHPNNANTMNDLTLKLLASPARYTAQVSPVCLASTNEALPSGL 120
DB 93 SDEENIQVLKIAKVFKNPKSILTVNNDITLKLATPARFSQTVSAVCLPSADDDFPAGT 152
QY 121 TCVTGTGGRISGVGNVTPARLQVVLPLVTNOCROYGARITDAMICAGSGASSCGD 180
DB 153 LCATGTGKTKYNANKTPDKLQQAALPLLSNAECKSGRRITDVMICAGASGVSSCMGD 212
QY 181 SGGPLVCQKAGNTWVLIGIVSGMGTKNICQAPAMVTRYSKFSTWYNQVAYN 231
DB 213 SGGPLVCQKAGANTLVIGVSGSDTCTSSPGVYARVTKLPWVQKILAAAN 263

RESULT 5
CTRB_BOVIN
```

ID CTRB\_BOVIN STANDARD; PRT; 245 AA.  
 AC P00767;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Chymotrypsinogen B (EC 3.4.21.1).  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE, DISULFIDE BONDS, AND ACTIVE SITE.  
 RX MEDLINE=68238908; PubMed=5649671;  
 RA Smillie L.B., Furka A., Nagabhushan N., Stevenson K.J., Parkes C.O.;  
 RT "Structure of chymotrypsinogen B compared with chymotrypsinogen A and  
 trypsinogen."; Biochem. J. 101:214-228(1966).  
 RL Nature 218:343-346(1968).  
 CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,  
 Phe-I-Xaa, Leu-I-Xaa.  
 CC -|- SUBCELLULAR LOCATION: Extracellular.  
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -|- DATABASE: NAME=Worthington enzyme manual;  
 WWW="http://www.worthington-biochem.com/manual/C/CHY.html".  
 DR PIR; A00953; KYBOB.  
 DR HSSP; P00766; IACB.  
 DR MEROPS; S01.152; -.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; TRYP\_SPC; 1.  
 DR PROSITE; PS00240; TRYP\_SIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYP\_SIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYP\_SIN\_SER; 1.  
 KW Hydrolyase; Serine protease; Digestion; Pancreas; Zymogen.  
 FT CHAIN 1 13  
 FT CHAIN 16 146 CHYMOTRYPSIN B, A CHAIN.  
 FT CHAIN 149 245 CHYMOTRYPSIN B, B CHAIN.  
 FT ACT\_SITE 57 57 CHYMOTRYPSIN B, C CHAIN.  
 FT ACT\_SITE 102 102 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 195 195 CHARGE RELAY SYSTEM.  
 FT DISULFID 1 122 CHARGE RELAY SYSTEM.  
 FT DISULFID 42 58  
 FT DISULFID 136 201  
 FT DISULFID 168 182  
 FT DISULFID 191 220  
 SQ SEQUENCE 245 AA; 25755 MW; 678016445FF5FEB5 CRC64;  
 Query Match 59.1%; Score 740.5; DB 1; Length 245;  
 Best Local Similarity 57.1%; Pred. No. 2.4e-59;  
 Matches 132; Conservative 32; Mismatches 66; Indels 1; Gaps 1;  
 QY 1 IYGENAVPGSPWQVSLQDNTGFHFCGSLISPNWVYTAACVTPGRHFVVLGEYDRS 60  
 DB 16 IYNGDAVPGSPWQVSLQDSTGFHFCGSLISDQWVYTAACGVTTSDVVVAGEFDQG 74  
 QY 61 SNAEPQVLSIARATHPPNWNANTMNDLTLKLASPARYTAQVSPVCLASTNEALPSGL 120  
 DB 75 LETEDQVQLKIGKVEKPKFSILTVRNDITLLKATPAQFSETVSACVLPSEDEFPGM 134  
 QY 121 TCVTGWGRISGVNVTTPARLQOVVLPVTVNQCRQYRGARTDAMICAGSGSCQGD 180  
 DB 135 LCATTGWGRTKYNALKTPDKLQOATLPVSNTPDCRYNGSRVTDVNICAGASGVSCMGD 194  
 QY 181 SGGLVPCQKNTWVLGIVSWCKTKNCNIOAPAMYTRVSKFTWVQVYMN 231  
 DB 195 SGGLVPCQKNGAWTLGIVSWGSGTCTSTTPAVYARVATLMPWQVETLAAN 245  
 RESULT 6  
 ID CTRB\_BOVIN STANDARD; PRT; 245 AA.

AC P00766;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Chymotrypsinogen A (EC 3.4.21.1).  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE, DISULFIDE BONDS, AND ACTIVE SITE.  
 RX MEDLINE=67181721; PubMed=5971783;  
 RA Brown J.R., Hartley B.S.;  
 RT "Location of disulphide bridges by diagonal paper electrophoresis.  
 The disulphide bridges of bovine chymotrypsinogen A."; Biochem. J. 101:214-228(1966).  
 RL Biochem. J. 101:214-228(1966).  
 RN [2]  
 RP REVISION TO 102.  
 RX MEDLINE=69106266; PubMed=5764436;  
 RA Blow D.M., Birktoft J.J., Hartley B.S.;  
 RT "Role of a buried acid group in the mechanism of action of  
 chymotrypsin."; Nature 221:337-340(1969).  
 RL Nature 221:337-340(1969).  
 RN [3]  
 RP PRELIMINARY SEQUENCE.  
 RA Hartley B.S.;  
 RT "Amino-acid sequence of bovine chymotrypsinogen-A."; Nature 201:1284-1287(1964).  
 RL Nature 201:1284-1287(1964).  
 RN [4]  
 RP SEQUENCE, AND DISULFIDE BONDS.  
 RX MEDLINE=67183948; PubMed=5972866;  
 RA Meloun B., Kluh I., Kostka V., Moravek L., Prusik Z., Vanacek J.,  
 Keil B., Sorm F.;  
 RT "Covalent structure of bovine chymotrypsinogen A."; Biochim. Biophys. Acta 130:543-546(1966).  
 RL Biochim. Biophys. Acta 130:543-546(1966).  
 RN [5]  
 RP ACTIVE SITE.  
 RX MEDLINE=67181723; PubMed=5971785;  
 RA Smillie L.B., Hartley B.S.;  
 RT "Histidine sequences in the active centres of some 'serine' proteinases."; Biochem. J. 101:232-241(1966).  
 RL Biochem. J. 101:232-241(1966).  
 RN [6]  
 RP X-RAY CRYSTALLOGRAPHY.  
 RX MEDLINE=72035052; PubMed=4399050;  
 RA Birktoft J.J., Blow D.M., Henderson R., Steitz T.A.;  
 RT "I. Serine proteinases. The structure of alpha-chymotrypsin."; Philos. Trans. R. Soc. Lond., B, Biol. Sci. 257:76-76(1970).  
 RL Philos. Trans. R. Soc. Lond., B, Biol. Sci. 257:76-76(1970).  
 RN [7]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF CHYMOTRYPSIN.  
 RX MEDLINE=70177557; PubMed=5442169;  
 RA Freer S.T., Kraut J., Robertus J.D., Wright H.T., Xuong N.H.;  
 RT "Chymotrypsinogen: 2.5-A crystal structure, comparison with alpha-chymotrypsin, and implications for zymogen activation."; Biochemistry 9:1997-2009(1970).  
 RL Biochemistry 9:1997-2009(1970).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF GAMMA-CHYMOTRYPSIN.  
 RX MEDLINE=82078042; PubMed=6914398;  
 RA Cohen G.H., Silverton E.W., Davies D.R.;  
 RT "Refined crystal structure of gamma-chymotrypsin at 1.9-A resolution. Comparison with other pancreatic serine proteases."; J. Mol. Biol. 148:449-479(1981).  
 RL J. Mol. Biol. 148:449-479(1981).  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (1.68 ANGSTROMS) OF ALPHA-CHYMOTRYPSIN.  
 RX MEDLINE=86011575; PubMed=4046030;  
 RA Tsukada H., Blow D.M.;  
 RT "Structure of alpha-chymotrypsin refined at 1.68-A resolution."; J. Mol. Biol. 184:703-711(1985).  
 RL J. Mol. Biol. 184:703-711(1985).  
 CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,  
 Phe-I-Xaa, Leu-I-Xaa.  
 CC -|- SUBCELLULAR LOCATION: Extracellular.  
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.







```

DR EMBL; L00119; AAA98780.1; JOINED.
DR EMBL; L00120; AAA98780.1; JOINED.
DR EMBL; L00121; AAA98780.1; JOINED.
DR EMBL; L00122; AAA98780.1; JOINED.
DR EMBL; L00123; AAA98780.1; JOINED.
DR PIR; A00961; ELRT2.
DR HSSP; P00772; IELG.
DR MEROPS; S01.155; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Hydrolase; Serine protease; Zymogen; Signal.
FT SIGNAL 1 16
FT PROPEP 17 30
FT CHAIN 31 271
FT ACT_SITE 75 75
FT ACT_SITE 123 123
FT ACT_SITE 218 218
FT DISULFID 60 76
FT DISULFID 157 224
FT DISULFID 188 204
FT DISULFID 214 245
SQ SEQUENCE 271 AA; 28885 MW; 125C783B857B71E3 CRC64;

Query Match 39.9%; Score 499; DB 1; Length 271;
Best Local Similarity 43.0%; Pred. No. 1.2e-37;
Matches 104; Conservative 37; Mismatches 89; Indels 12; Gaps 8;

QY 1 IVNGENAVPGSMWQVSLQ---DNTGFHFCGSLISPNWVTAHCOVTPGRHFVVLGEY 57
DB 31 VVGQASNSPWPQVSLQYLSGKWHHTCGGSLVANNVLTAAHCISNRYRVLGRH 90
QY 58 DRSSNAEPQVLSIARAIHPNNANTMN--NDLTLLKLASPARYTAQVSPVCLASTNEA 115
DB 91 SLSTSESGSLAVQVSLVHVKWNAQKLSNGNDIALVKLASPVALTSKIOTACLPAGTI 150
QY 116 LPSEGLTCVTGWRISGVGNVTPARLQVPLVTVNQC--RQYWGARITDMICAGSGG 173
DB 151 LPNNYPCYVTGWRIGLTNG--ATPDVLQGGRLVVDYATCSSASWWSGSSVKTNNMVCAGDG 209
QY 174 -ASSCGDGGPLVCCQGN--TWVLIGVSWG-TKNCNI-QAPAMYTRVSKFSTWVNOVMA 229
DB 210 VTSNCNGDGGPLNCQASNGQWQVHGIVSGFTLGCNYPKPSVFTRVSNYIDWINSVIA 269
QY 230 YN 231
DB 270 KN 271

RESULT 11
CAC3_BOVIN STANDARD; PRT; 253 AA.
ID CAC3_BOVIN
AC P05805;
DT 01-NOV-1988 (Rel. 09, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Proprioteinase E precursor (Procarboxypeptidase A complex component
DE III) (Procarboxypeptidase A-S6 subunit III) (PROCPA-S6 III).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE OF 1-25.
RX MEDLINE=91099520; PubMed=2269366;
RA Pascual R., Vendrell J., Aviles F.X., Bonicel J., Wicker C.,
RA Puigserver A.;

```

```

RT "Autolysis of proproteinase E in bovine procarboxypeptidase A ternary
RT complex gives rise to subunit III.";
RL FEBS Lett. 277:37-41(1990).
RN [2]
RP SEQUENCE OF 14-253, AND DISULFIDE BONDS.
RX MEDLINE=86220198; PubMed=3519215;
RA Venot N., Sciaky M., Puigserver A., Desnuelle P., Laurent G.;
RT "Amino acid sequence and disulfide bridges of subunit III, a
RT defective endopeptidase present in the bovine pancreatic 6 S
RT procarboxypeptidase A complex.";
RL Eur. J. Biochem. 157:91-99(1986).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).
RX MEDLINE=94222022; PubMed=8168476;
RA Pignol D., Gaboriaud C., Michon T., Kerfelec B., Chapus C.;
RA Fontecilla-Camps J.C.;
RT "Crystal structure of bovine procarboxypeptidase A-S6 subunit III, a
RT highly structured truncated zymogen E.";
RL EMBO J. 13:1763-1771(1994).
CC -!- FUNCTION: DEFECTIVE ELASTASE-LIKE SERINE PROTEASE. DOES NOT SEEM
CC TO HAVE A PROTEASE ACTIVITY. ITS LIKELY FUNCTION IS TO PROTECT
CC PROCARBOXYPEPTIDASE A AGAINST DENATURATION IN THE ACIDIC
CC ENVIRONMENT OF THE RUMINANT DUODENUM.
CC -!- SUBUNIT: HETEROPRIMER OF SUBUNIT III; CARBOXYPEPTIDASE A AND
CC CHYMOTRYPSINOGEN C.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: PANCREAS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR PIR; A25065; CPBOA3.
DR PDB; 1FON; 14-OCT-96.
DR MEROPS; S01.983; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Serine protease homolog; Pancreas; Digestion; 3D-structure.
KW PROPEP 1 11
FT ACTIVATION PEPTIDE.
FT CHAIN 12 253
FT DISULFID 41 57
FT DISULFID 100 103
FT DISULFID 140 206
FT DISULFID 171 187
FT DISULFID 196 227
SQ SEQUENCE 253 AA; 27337 MW; 24663724D8AE409C CRC64;

Query Match 39.7%; Score 497.5; DB 1; Length 253;
Best Local Similarity 43.6%; Pred. No. 1.5e-37;
Matches 106; Conservative 36; Mismatches 88; Indels 13; Gaps 8;

QY 1 IVNGENAVPGSMWQVSLQ---DNTGFHFCGSLISPNWVTAHCOVTPGRHFVVLGEY 57
DB 12 VVGQASNSPWPQVSLQYLSGKWHHTCGGSLVANNVLTAAHCISNRYRVLGRH 71
QY 58 DRSSNAEPQVLSIARAIHPNNANTMN--NDLTLLKLASPARYTAQVSPVCLASTN 113
DB 72 DRSVLESGEQVPIINAGDLFVHPVPLWNSNCVACGNDIALVKLRSQALGDKVQLANLPPAG 131
QY 114 EALPSGLTCVTGWRISGVGNVTPARLQVPLVTVNQCRO--YWGARITDMICAGG 171
DB 132 DILNEAPCVISGWGRLY-TGGPLPKLQALPLVDVYEHCSQMDWNGITVTKTMCVAGG 190
QY 172 SGASSCGDGGPLVCCQGN--TWVLIGVSWG-TKNCNI-QAPAMYTRVSKFSTWVNOVMA 228
DB 191 DTRSGCNGDGGPLNCPAADGSMQVHCVTSFVSAFGCNTIKKPTVTRVSAFIDWIDETI 250
QY 229 AYN 231
DB 251 ASN 253

```

DR	PRINTS; PRO0722; CHYMOTRYPSIN.
DR	SMART; SMO0020; TRYF_SPC; 1.
DR	PROSITE; PS0240; TRYPsin_DOM; 1.
DR	PROSITE; PS00134; TRYPsin_HIS; 1.
DR	PROSITE; PS00135; TRYPsin_SER; 1.
KW	Hydrolase; Serine protease; Zymogen; Signal.
FT	SIGNAL 1 16
FT	PROPEP 17 28 ACTIVATION PEPTIDE.
FT	CHAIN 29 269 ELASTASE 2A.
FT	DISULFID 58 74 BY SIMILARITY.
FT	DISULFID 155 222 BY SIMILARITY.
FT	DISULFID 186 202 BY SIMILARITY.
FT	DISULFID 212 243 BY SIMILARITY.
FT	ACT_SITE 73 73 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 121 121 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 216 216 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	CONFLICT 202 202 C -> V (IN REF. 3).
SQ	SEQUENCE 269 AA; 28888 MW; A2E05143EFF4987C CRC64;
Query Match 38.8%; Score 486; DB 1; Length 269;	
Best Local Similarity 43.0%; Pred. No. 1.7e-36;	
Matches 104; Conservative 41; Mismatches 85; Indels 12; Gaps	
QY	1 IVGENAVPGSPWPQVSLQNTG---FHFCGGSLIPNWWVTAAHCOVTPGRHFVVLEGEY 57 :   :          : :         :   :    :   :  :
Db	29 VVGEEARPNSWPQVSLQYSNGKWKYHTCGGLIANSWVLTAAHGISSTRTYRVGLGRH 88 :   :          : :         :   :    :   :  :
QY	58 DRSSNAEPQVLSLSTARITHPNKNATMN--NDLTLLKLASPARYAQAQVPVCLASTNEA 115 :   :    :   :  :   :          :   :    :   :  :
Db	89 NLYVAESGLVASVSKITVHHKDWSNQISKGDIALKLANVSLTDKIQLACLPPAGTI 148    :          :   :          :   :          :   :          :
QY	116 LPSCGLTCVTGWGRISGVGNVTPARQQVPLPLVTVNQCRC--YWGARTDAMICAGSG 173    :          :   :          :   :          :   :          :
Db	149 LPNNYPCYVTGWRLQTNGAV-PDVLOOGRLLVVDVATCSSAWGSSVKTSICAGDG 207    :          :   :          :   :          :   :          :
QY	174 A-SQCSDGSGPLVCQKGN-TWVLIGIVSWGTK-NCN-IQAPAMYTRVSKFSTWINOVMA 229    :          :   :          :   :          :   :          :
Db	208 VISSCNGDSGPLNCQASDRGWQVHGIVSFGSLGCNYHKPSVFTRVSNIWDINSVIA 257    :          :   :          :   :          :   :          :
QY	230 YN 231
Db	268 NN 269
RESULT 13	
EL3B_HUMAN	STANDARD; PRT; 270 AA.
ID	EL3B_HUMAN STANDARD; PRT; 270 AA.
AC	P08861; P11423;
DT	01-NOV-1988 (Rel. 09, Created)
DD	16-OCT-2001 (Rel. 40, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Elastase IIIB precursor (EC 3.4.21.70) (Protease E).
DN	ELA3B.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Pancreas;
RX	MEDLINE=88087253; PubMed=2826474;
RA	Tani T., Ohsumi J., Mita K., Takiguchi Y.;
RT	"Identification of a novel class of elastase isozyme, human
RT	pancreatic elastase III, by cDNA and genomic gene cloning.";
RL	J. Biol. Chem. 263:1231-1239(1988).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Pancreas;
RA	Strausberg R.;
RL	Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
RN	[3]
RP	SEQUENCE OF 4-270 FROM N.A.
RC	TISSUE=Pancreas;

```

FT CONFLICT 129 131 MISSING (IN REF. 5).
FT CONFLICT 164 164 R -> P (IN REF. 3).
SQ SEQUENCE 270 AA; 29293 MW; B14BE0AAD369SAFE CRC64;

Query Match 38.5%; Score 481.5; DB 1; Length 270;
Best Local Similarity 42.7%; Pred. No. 4.3e-36;
Matches 103; Conservative 39; Mismatches 86; Indels 13; Gaps 9;

QY 1 IVGCENAVPGSWPQVWSLO--DNTG--FHFGCGSLISPNWVVTAAHCOVTPCRHFHVVLGEY 57
   :|||:|||:|||:|||:::|:|||:|||:|||:::|||||
Db 29 VVNGEDAVPYSPWPQVWSLQYEKSGSFHTCGGSLIADPWVVTAGHCILSSWTQVVLGEY 88
   ||:|:||:|||||:::|:|||:|||:|||:::|||||
QY 58 DRGSNAEPVQVLSI--ARAITHPNWNANTM--NNDLTLLKLASPARYTQAQVPVCLASTN 113
   ||:|:||:|||||:::|:|||:|||:|||:::|||||
Db 89 DRAVKEGPEQVIPINSGLFVPHLPNNSCVACGNDIALIKLSRSAQLGDVAQLASLPAG 148
   ||:|:||:|||||:::|:|||:|||:|||:::|||||
QY 114 EALPSGLTCTVTTGWGRISGVGNVTPARLQQVPLTVTNOCRQ--YWGARITDMICAGG 171
   :||:|:|||::|:|||::|:|||::|:|||::|:|||::|:|||::|:|||
Db 149 DILPFNETPCYITWGRLTYTGPL-PDKLQALLPVVDYEHCSRNNWNGSSVKKTMCAGG 207
   ||:|:||:|||||:::|:|||:|||:|||:::|||||
QY 172 SGASSCCGGDGGLPVC-QKGTNWLVIGIVSW-GTKNCNI-QAPAMYTRVSKFSWIQVNM 228
   ||:|:||:|||||:::|:|||:|||:|||:::|||||
Db 208 DIRSGCMGDSGPLNCPTEGGMQVHGTVSVSFAFGCNTRKPFTVRVSAFIDWIEETI 267
   ||:|:||:|||||:::|:|||:|||:|||:::|||||
QY 229 A 229
   |
Db 268 A 268

RESULT 14
EL2_BOVIN ID EL2_BOVIN STANDARD; PRT; 269 AA.
AC Q29461;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
Elastase 2 precursor (EC 3.4.21.71).
GN ELA2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
NCBI_TaxID=9913;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Pancreas;
RC MEDLINE=98079203; PubMed=9418008;
RX Gestin M., le Huerou-Luron I., Wicker-Planquart C., le Drean G.,
RA Chaix J.C., Puigserver A., Guilloteau P.;
RT "Bovine pancreatic preelastases I and II: comparison of nucleotide
RT and amino acid sequences and tissue specific expression.";
RL Comp. Biochem. Physiol. 118B:181-187(1997).
CC -|- FUNCTION: ACTS UPON ELASTIN.
CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Leu-|-Xaa, Met-|-Xaa
CC and Phe-|-Xaa. Hydrolyzes elastin.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- TISSUE SPECIFICITY: PANCREAS.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; X97635; CAA66231.1; -
CC HSSP; P00766; LCHG.
CC MEROPS; S01.155; -.
DR InterPro; IPRO01314; Chymotrypsin.
DR InterPro; IPRO01254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.

```



[illegible]



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:14:28 ; Search time 14.6087 Seconds  
(without alignments)  
1520.126 Million cell updates/sec

Title: US-09-856-319B-4\_COPY\_34\_264

Perfect score: 1252

Sequence: 1 IVNGENAVPGSWPQVSLQD.....AMTVRSKFTWVNOVMAYN 231

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_73.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1119	89.4	264	2 I38136	chymotrypsin-like
2	757.5	60.5	263	2 A21195	chymotrypsin (EC 3
3	750.5	59.9	263	2 S47537	chymotrypsin (EC 3
4	750.5	59.9	263	2 A31299	chymotrypsin (EC 3
5	740.5	59.1	245	1 KYBOB	chymotrypsin (EC 3
6	727.5	58.1	245	1 KYBOA	chymotrypsin (EC 3
7	715.5	57.1	263	1 KYRTB	chymotrypsin (EC 3
8	714.5	57.1	244	2 S72219	chymotrypsin B - A
9	503	40.2	269	2 A26823	pancreatic elastase
10	499	39.9	271	1 ELRT2	pancreatic elastase
11	487	38.9	1524	2 T30337	polyprotein - Afri
12	486	38.8	269	2 B26823	pancreatic elastase
13	485.5	38.8	240	1 CPBOA3	procarboxypeptidase
14	481.5	38.5	270	2 B29934	pancreatic elastase
15	475	37.9	271	2 A25528	pancreatic elastase
16	464.5	37.1	810	1 PLHU	plasmin (EC 3.4.21
17	462	36.9	812	1 PLMS	plasmin (EC 3.4.21
18	460.5	36.8	270	2 A29934	pancreatic elastase
19	455.5	36.4	270	2 S56160	mast cell tryptase
20	452.5	36.1	460	2 B61545	plasmin (EC 3.4.21
21	451.5	36.1	810	2 B30848	plasmin (EC 3.4.21
22	449.5	35.9	790	1 PLPG	plasmin (EC 3.4.21
23	449	35.9	269	2 B26823	pancreatic elastase
24	448.5	35.8	343	1 A57014	proctasin (EC 3.4
25	447.5	35.7	268	2 S68826	pancreatic elastase
26	447.5	35.7	268	2 S68826	pancreatic elastase
27	447.5	35.7	273	2 A47246	tryptase (EC 3.4.2
28	444	35.5	231	1 TRPGTR	trypsin (EC 3.4.21
29	441	35.2	126	2 A23473	chymotrypsin-like

30	440	35.1	246	1 TRRT1	trypsin (EC 3.4.21
31	439.5	35.1	274	2 JC4171	tryptase (EC 3.4.2
32	437.5	34.9	455	2 A61545	plasmin (EC 3.4.21
33	436	34.8	246	1 TRRT2	trypsin (EC 3.4.21
34	436	34.8	246	1 TRDGC	trypsin (EC 3.4.21
35	433	34.6	247	1 TRDGC	trypsin (EC 3.4.21
36	432	34.5	248	2 S55067	trypsin (EC 3.4.21
37	430.5	34.4	266	1 ELRT1	pancreatic elastase
38	430.5	34.4	276	2 A38654	mast cell proteina
39	430	34.3	238	2 S31779	trypsin (EC 3.4.21
40	429	34.3	1019	1 A56318	enteropeptidase (E
41	428.5	34.2	275	2 A32410	tryptase (EC 3.4.2
42	428	34.2	810	2 I46260	plasmin (EC 3.4.21
43	427.5	34.1	258	2 S70439	pancreatic elastase
44	427.5	34.1	267	4 A56615	probable pancreati
45	427.5	34.1	275	2 C35863	tryptase (EC 3.4.2

#### ALIGNMENTS

##### RESULT 1

I38136

chymotrypsin-like proteinase (EC 3.4.21.-) CTRL-1 - human

C:Species: Homo sapiens (man)

C:Date: 06-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change 22-Jun-1999

C:Accession: I38136

R:Larsen, F.; Solheim, J.; Kristensen, T.; Kolsto, A.B.; Prydz, H.

Hum. Mol. Genet. 2, 1589-1595, 1993

A:Title: A tight cluster of five unrelated human genes on chromosome 16q22.1.

A:Reference number: I38135; MUID:94093544; PMID:8268911

A:Accession: I38136

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-264 <RES>

A:Cross-references: EMBL:X71874; NID:g406226; PIDN:CAA50710.1; PID:g406228

C:Genetics:

A:Gene: GDB:CTRL

A:Cross-references: GDB:204061

A:Map position: 16q22.1-16q22.1

A:Introns: 18/1; 52/3; 79/2; 106/3; 167/1; 211/3

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; serine proteinase

F:34-257/Domain: trypsin homology <TRY>

F:75.121.214/Active site: His, Asp, Ser #status predicted

Query Match

Best Local Similarity 89.4%; Score 1119; DB 2; Length 264;

Matches 199; Conservative 22; Mismatches 10; Indels 0; Gaps 0;

QY 1 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVTAACOVTPGRHFVVLGYDRS 60

Db 34 IVNGENAVPGSWPQVSLQDSSGFHFCGSLISQSWVTAACNVSFGRHFVVLGYDRS 93

QY 61 SNAEPQVLSIARATHPNWNTANNLTLKSLASPARYTAQVSPVCLASNEALPSGL 120

Db 94 SNAEPQVLSIARATHPNWNTANNLTLKSLASPARYTAQVSPVCLASNEALPSGL 153

QY 121 TCVTTGWRISGVNVTAPLQOVPLVTVNQCRQYWGARTDAMICAGGSSCCQGD 180

Db 154 TCVTTGWRISGVNVTAPLQOVPLVTVNQCRQYWGARTDAMICAGGSSCCQGD 213

QY 181 SGGPLVCOKNTWLLIGIVSWGFKNCNICAPAMTVRSKFTWVNOVMAYN 231

Db 214 SGGPLVCOKNTWLLIGIVSWGFKNCNICAPAMTVRSKFTWVNOVMAYN 264

##### RESULT 2

A21195

chymotrypsin (EC 3.4.21.1) 2 precursor - dog

C:Species: Canis lupus familiaris (dog)

C:Date: 27-Jul-1990 #sequence\_revision 27-Jul-1990 #text\_change 22-Jun-1999

C:Accession: A21195



F.16-238/Domain: trypsin homology <TRY>  
F.1-122,42-58,136-201,168-182,191-220/Disulfide bonds: #status experimental  
F.57,102,195/Active site: His, Asp, Ser #status experimental

Query Match	58.1%	Score 727.5;	DB 1;	Length 245;
Best Local Similarity	56.7%;	Pred. No. 6.4e-56;		
Matches 131;	Conservative 32;	Mismatches 67;	Indels 1;	Gaps

```

QY      1  IVNGENAVPGSNPWQVSLQDNTGFHFCGSLISPNWVYTAACHQVTPGRHFVWLGEYDRS 60
        ||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      16  IVNGEAVPGSNPWQVSLQDKTGFHFCGSLINWVYTAARCGVTTS-DVVVAGEFDQG 74
        ||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY      61  SNAEPVOVLRIARTHPNWNANTMNDLTLLKLPARYTAQVSPVCLASTNEALPSGL 120
        |:| :| | | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      75  SSSEKIQLKIAKFVKNSKYLNITNDITLLKLSTAASFQSVAVCLPSASDDFAAGT 134
        ||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY      121 TCVTTGWGRISGVGNVTPARLOOQVPLVTVNQCYNGARITDAMICAGSGASSCOGD 180
        ||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      135 TCVTTGWGTRTYNANTPDLQAQSLPLSNTNKCKYWGTKIKIDAMICAGASGVSSCMGD 194
        ||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY      181 SGGPLVCQRGNTFWLIGISWGTKNICIQAPAMYTRVSKFTSWINQMAYN 231
        ||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      195 SGGPLVCKKGAWTLVGIVSWGSSCTSTPCGYARVATLVNWVQQTLAAN 245
        ||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 7
KYRTH
Chymotrypsin (EC 3.4.21.1) B precursor - rat
N:Alternate names: chymotrypsinogen B
C:Species: Rattus norvegicus (Norway rat)
C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 18-Jun-1999
C:Caccession: A22658
J:R.Bell, G.I.; Quinto, C.; Quiroga, M.; Valenzuela, P.; Craik, C.S.; Rutter, J. Biol. Chem. 259, 14265-14270, 1984
A:Title: Isolation and sequence of a rat chymotrypsin B gene.
A:Reference number: A22658; MUID:85054881; PMID:6209274
A:Accession: A22658
A:Molecule type: DNA
A:Residues: 1-263 <BEL>
A:Cross-references: GB:K02298; NID:g203653; PIDN:AAA98732.1; PID:g203654
C:Genetics:
A:Introns: 18/1; 52/3; 79/2; 105/3; 166/1; 210/3
C:Superfamily: tryptase; trypsin homology
C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase
F.1-18/Domain: signal sequence #status predicted <SIG>
F.19-33/Domain: propeptide #status predicted <PRO>
F.34-263/product: chymotrypsin B #status predicted <MAT>
F.34-256/Domain: trypsin homology <TRY>
F.75,120,213/Active site: His, Asp, Ser #status predicted
```

Query Match	57.1%	Score 715.5;	DB 1;	Length 263;
Best Local Similarity	55.4%;	Pred. No. 7.6e-55;		
Matches 128;	Conservative 35;	Mismatches 67;	Indels 1;	Gaps

```

QY      1  IVNGENAVPGSNPWQVSLQDNTGFHFCGSLISPNWVYTAACHQVTPGRHFVWLGEYDRS 60
        ||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      34  IVNGDAIPGSNPWQVSLQDKTGFHFCGSLISDWVYTAARCGVKTS-DVVVAGEFDQG 92
        ||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY      61  SNAEPVOVLRIARTHPNWNANTMNDLTLLKLPARYTAQVSPVCLASTNEALPSGL 120
        |:| :| | | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      93  SDENIOLVLIQAQVFNKPFNFRTNRDITLLKLATPAQFSSETSAVCLPNVDODFFPGT 152
        ||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY      121 TCVTTGWGRISGVGNVTPARLOOQVPLVTVNQCYNGARITDAMICAGSGASSCOGD 180
        ||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      153 VCATTGWGKTNYALKTPBKLQAAPIVSEADCKKSWGSKITDVNTCAGASGVSSCMGD 212
        ||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY      181 SGGPLVCQRGNTFWLIGISWGTKNICIQAPAMYTRVSKFTSWINQMAYN 231
        ||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      213 SGGPLVCQRGVTLAGIVSWGSGVCSSTPAPYSRVATLMPPVQOILEAN 263
        ||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 8
S72219
```

QY	1	IVGENAVPGSWPQWSLQ--DNTG--FHFCGSLISPNVWVTAAHQCVTPGRHFVVVLGEY	57
Db	29	VVGEDARPNSWPQWSLQYDSSGQWRHTCGGTLPDQSWLTAACHICSSSRTYRVVLGRH	88
QY	58	DRSNAEPQVLSIARIAITHPNKNAITMN--NDLTLLKLSAPRYTAQVSPVCLASTNEA	115
Db	89	SLSTNEPGLAVKSLKVLWQDWSNQLSGNDIALKLASPSVLTDKLQLGCLPAAGTY	148
QY	116	LPSELCTVTGWGRISGVGNVTTPARLQVVLPLVTVYQCRCQ--YMGARITDAMICAGSG	173

A;RESIDUES: 1-1324 <IAN>  
A;Cross-references: EMBL:U81290; NID:q2981640; PID:q2981641; PIDN:AAC24717.1

Query Match 38.9%; Score 487; DB 2; Length 1524;  
Best Local Similarity 39.5%; Pred. No. 4.2e-34;  
Matches 94; Conservative 49; Mismatches 75; Indels 20; Gaps 6;

Qy 1 IVGENAVPGSWPQVSLQDNTG---FHFCGSLISPNWVYTAACHQVTPGRHFVVLGEYDRS 60  
Db 1295 VVGQQAAPRSWPLVLSIQNSKRHYCGGIIITNKWILTAACHCEVKINLHRVYVVGHTDLT 1354  
Qy 61 --SNAEPQVLSIARATHPNWNTMNDLTLLKLASPARYTAQVSPVCLASTNEALPS 118  
Db 1355 EVQNEHAKYVNSHVELYMP--GSPPRNDLLLELDTPLLNNSVAVIC-----LPD 1405  
Qy 119 GLT-----CVTTGWRISGVNVT--PARLQVQVPLVTVNQCROXYWGARITDAMICAG 170  
Db 1406 DVTDTQAECLVAGWG-VTDVGMSLPTKLQQAQVPIVSTFKCKDYWSDVTDNNICAG 1464  
Qy 171 GSGASSCGDSGGLPVCKQNTWVLIGIVSWGTKNCNTQAPAMTRYKSTFWINQVM 228  
Db 1465 KAGASSCMGDSGGLPCKREDRYVLGVVSWGSGKCDVKAPSVYTLTSAFMDWISQHM 1522

RESULT 12  
B26823  
pancreatic elastase II (EC 3.4.21.71) A precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 16-Aug-1988 #sequence\_revision 16-Aug-1988 #text\_change 20-Jun-2000  
C:Accession: B26823; A27432; A41431; S34491  
R:Kawashima, I.; Tani, T.; Shimoda, K.; Takiguchi, Y.  
DNA 6, 163-172, 1987  
A:Title: Characterization of pancreatic elastase II cDNAs: two elastase II mRNAs are expressed in the pancreas of the rat  
A:Reference number: A90958; MUID:87217962; PMID:3646943  
A:Accession: B26823  
A:Molecule type: mRNA  
A:Residues: 1-269 <KAW>  
A:Cross-references: GB:M16652; NID:g182057; PIDN:AAA52380.1; PID:g182058  
R:Fletcher, T.S.; Shen, W.F.; Largman, C.  
Biochemistry 26, 7256-7261, 1987  
A:Title: Primary structure of human pancreatic elastase 2 determined by sequence analysis  
A:Reference number: A27432; MUID:88107669; PMID:3427074  
A:Accession: A27432  
A:Molecule type: mRNA  
A:Residues: 1-269 <FLE>  
A:Cross-references: GB:M16631; NID:g182022; PIDN:AAA52374.1; PID:g182023  
R:Shiraga, Y.; Yoshida, H.; Matsuki, S.; Takemura, K.; Ikeda, N.; Shimada, Y.; Ozawa, T.  
J. Biochem. 102, 1553-1563, 1987  
A:Title: Molecular cloning and expression in Escherichia coli of a cDNA encoding human pancreatic elastase II  
A:Reference number: A41431; MUID:88198076; PMID:2834346  
A:Accession: A41431  
A:Molecule type: mRNA  
A:Residues: 1-201, 'V', 203-269 <SHI>  
A:Cross-references: GB:D00236; NID:g219619; PIDN:BAA00165.1; PID:g219620  
A:Note: the authors translated the codon GTG for residue 202 as Cys  
R:Moulard, M.; Michon, T.; Kerfelec, B.; Chapus, C.  
FEBS Lett. 261, 179-183, 1990  
A:Title: Further studies on the human pancreatic binary complexes involving procarboxypeptidase Y  
A:Reference number: S08253; MUID:90169111; PMID:2307232  
A:Accession: S34491  
A:Molecule type: protein  
A:Residues: 'X', 18-50 <MOU>  
C:Genetics:  
A:Gene: GDB:ELAI  
A:Cross-references: GDB:I19866; OMIM:130120  
A:Map position: 12pter-12qter  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; pancreas; serine proteinase  
F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-28/Domain: propeptide #status predicted <PRO>  
F:29-269/Product: pancreatic elastase IIA #status predicted <MAT>  
F:29-262/Domain: trypsin homology <TRY>  
F:73,121,216/Active site: His, Asp, Ser #status predicted

Query Match 38.8%; Score 486; DB 2; Length 269;  
Best Local Similarity 43.2%; Pred. No. 6.9e-35;  
Matches 104; Conservative 36; Mismatches 88; Indels 13; Gaps 8;

Qy 3 NGENAVPGSWPQVSLQ---DNTGFHFCGSLISPNWVYTAACHQVTPGRHFVVLGEYDR 59  
Db 1 DGEDAVPYSWNQVSLQYKEKGAFHHTCGSLIAPDWVYTAGHCISTRTYQVVLGEYDR 60  
Qy 60 S--SNAEPQVLSIARATHPNWNTM--NNDLTLKLASPARYTAQVSPVCLASTNEA 115  
Db 61 SVLEGEQVPIINAGDLFVHPLWNSNCVACGNDIALVKLSRAQLGDKVQLANLPPAGDI 120  
Qy 116 LPSGLTCVTTGWRISGVNVT--PARLQVQVPLVTVNQCRO--YWGARITDAMICAGSGG 173  
Db 121 LPNEAPCYISGWRGLY-TGGPLPDKLQALLPVVDYEHCSQDWNGITVKKTMVCAAGDT 179  
Qy 174 ASSCGDSGGPLVCOKGN-TWVLIGIVSW-GTKNCN-IOAPAMTRYKSTFWINQVMAY 230  
Db 180 RSGCNGDSGGPLNCPPADGSGQVHGVTSFVSFAFGCNTIKKPTVTRVSAFIDWIDETIAS 239  
Qy 231 N 231  
Db 240 N 240

RESULT 14  
B29934

Best Local Similarity 43.0%; Pred. No. 7.1e-35;  
Matches 104; Conservative 41; Mismatches 85; Indels 12; Gaps 8;

Qy 1 IVGENAVPGSWPQVSLQDNTG---FHFCGSLISPNWVYTAACHQVTPGRHFVVLGEY 57  
Db 29 VVGEEARPNSWPQVSLQYSSNGKWYHTCGSLTANSWILTAACHCISSTRYRVGLGRH 88  
Qy 58 DRSSNAEPQVLSIARATHPNWNTM--NNDLTLKLASPARYTAQVSPVCLASTNEA 115  
Db 89 NLYVAESGSLAVSVKIVVHKDWSNQLSKGNDIALKLANPVSITDKIQLACLPPAGTI 148  
Qy 116 LPSGLTCVTTGWRISGVNVT--PARLQVQVPLVTVNQCRO--YWGARITDAMICAGSGG 173  
Db 149 LPNNYPCVVTGWRLOTNGAV-PDVLOQGRLLVVDYATCSSSAMWSSVKTSMICAGG 207  
Qy 174 A-SSCGDSGGPLVCOKGN-TWVLIGIVSWGK-NCN-IOAPAMTRYKSTFWINQVMA 229  
Db 208 VISSCNGDSGGPLNCQASDGRMVHGIYVSGRLGCTNYHKPSVTRVSNVDMINSVIA 267  
Qy 230 YN 231  
Db 268 NN 269

RESULT 13  
CPBOA3  
procarboxypeptidase A complex component III - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 18-Jul-1997  
C:Accession: A25065  
R:Venot, N.; Sciaky, M.; Puigserver, A.; Desnuelle, P.; Laurent, G.  
Eur. J. Biochem. 157, 91-99, 1986  
A:Title: Amino acid sequence and disulfide bridges of subunit III, a defective endopeptidase against denaturation in the acidic environment of the ruminant duodenum.  
A:Reference number: A25065; MUID:86220198; PMID:3519215  
A:Accession: A25065  
A:Molecule type: protein  
A:Residues: 1-240 <VRN>  
C:Comment: This protein is found in the pancreatic juice of ruminants, where it const-  
1) and chymotrypsin C (EC 3.4.21.2). It is inactive towards the usual substrates of  
ypeptidase against denaturation in the acidic environment of the ruminant duodenum.  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: pancreas  
F:1-233/Domain: trypsin homology <TRY>  
F:28-44, 87-90, 127-193, 158-174, 183-214/Disulfide bonds: #status experimental  
F:43,93,187/Active site: His, Asp, Ser #status predicted

Query Match 38.8%; Score 485.5; DB 1; Length 240;  
Best Local Similarity 43.2%; Pred. No. 6.9e-35;  
Matches 104; Conservative 36; Mismatches 88; Indels 13; Gaps 8;

Qy 3 NGENAVPGSWPQVSLQ---DNTGFHFCGSLISPNWVYTAACHQVTPGRHFVVLGEYDR 59  
Db 1 DGEDAVPYSWNQVSLQYKEKGAFHHTCGSLIAPDWVYTAGHCISTRTYQVVLGEYDR 60  
Qy 60 S--SNAEPQVLSIARATHPNWNTM--NNDLTLKLASPARYTAQVSPVCLASTNEA 115  
Db 61 SVLEGEQVPIINAGDLFVHPLWNSNCVACGNDIALVKLSRAQLGDKVQLANLPPAGDI 120  
Qy 116 LPSGLTCVTTGWRISGVNVT--PARLQVQVPLVTVNQCRO--YWGARITDAMICAGSGG 173  
Db 121 LPNEAPCYISGWRGLY-TGGPLPDKLQALLPVVDYEHCSQDWNGITVKKTMVCAAGDT 179  
Qy 174 ASSCGDSGGPLVCOKGN-TWVLIGIVSW-GTKNCN-IOAPAMTRYKSTFWINQVMAY 230  
Db 180 RSGCNGDSGGPLNCPPADGSGQVHGVTSFVSFAFGCNTIKKPTVTRVSAFIDWIDETIAS 239  
Qy 231 N 231  
Db 240 N 240

RESULT 14  
B29934

pancreatic elastase (EC 3.4.21.36) IIIB precursor - human  
N:Alternate names: 35K glycoprotein; pancreatic protein P35; protein G32; proteinase E  
C:Species: Homo sapiens (man)  
C:Date: 15-Dec-1988 #sequence\_revision 15-Dec-1988 #text\_change 22-Jun-1999  
C:Accession: B29934; A90516; A27206; A33257; A28932; S04490; S04490  
R:Tani, T.; Ohsumi, J.; Mita, K.; Takiguchi, Y.  
J. Biol. Chem. 263, 1231-1239, 1988  
A:Title: Identification of a novel class of elastase isozyme, human pancreatic elastase  
A:Reference number: A92664; MUID:88087253; PMID:2826474  
A:Accession: B29934  
A:Molecule type: mRNA  
A:Residues: 1-270 <TAN>  
A:Cross-references: GB:M18692; NID:9607029; PIDN:AAA58454.1; PID:g182035  
R:Shen, W.; Fletcher, T.S.; Largman, C.  
Biochemistry 26, 3447-3452, 1987  
A:Title: Primary structure of human pancreatic protease E determined by sequence analysis  
A:Reference number: A90516; MUID:88000545; PMID:3477287  
A:Accession: A90516  
A:Molecule type: mRNA  
A:Residues: 'G', 5-63, 'G', 65-78, 'W', 80-118, 'G', 120-163, 'P', 165-270 <SHE>  
R:Fletcher, T.S.  
Submitted to GenBank, August 1987  
A:Reference number: A94507  
A:Contents: revision to residue 119  
A:Accession: A27206  
A:Molecule type: DNA  
A:Residues: 'G', 5-63, 'G', 65-78, 'W', 80-118, 'G', 120-163, 'P', 165-270 <PLE>  
R:Aviles, F.X.; Pascual, R.; Salva, M.; Bonicel, J.; Puigserver, A.  
Biochem. Biophys. Res. Commun. 163, 1191-1196, 1989  
A:Title: Generation of a subunit III-like protein by autolysis of human and porcine pro-  
A:Reference number: A33257; MUID:89392022; PMID:2675835  
A:Accession: A33257  
A:Molecule type: protein  
A:Residues: 18-57 <AVI>  
R:Guy-Crotte, O.; Barthe, C.; Basso, D.; Rournet, B.; Figarella, C.  
Biochem. Biophys. Res. Commun. 156, 318-322, 1988  
A:Title: Characterization of two glycoproteins of human pancreatic juice: P35, a truncated  
A:Reference number: A28932; MUID:89025862; PMID:3178837  
A:Accession: A28932  
A:Molecule type: protein  
A:Residues: 'X', 32-52, 'X', 54-55, 'XXX', 59-63 <GUY>  
R:Moulard, M.; Kerfelec, B.; Mallet, B.; Chapus, C.  
FEBS Lett. 250, 166-170, 1989  
A:Title: Identification of a procarboxypeptidase A-truncated protease E binary complex  
A:Reference number: S04999; MUID:89325560; PMID:2753124  
A:Accession: S04999  
A:Molecule type: protein  
A:Residues: 31-50 <MOU>  
R:Wendorf, P.; Geyer, R.; Sziegoleit, A.; Linder, D.  
FEBS Lett. 249, 275-278, 1989  
A:Title: Localization and characterization of the glycosylation site of human pancreatic  
A:Reference number: S04490; MUID:89289996; PMID:2737288  
A:Accession: S04490  
A:Molecule type: protein  
A:Residues: 94-128;132-164 <WEN>  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: glycoprotein; hydrolase; serine proteinase; zymogen  
F:117/Domain: signal sequence #status predicted <SIG>  
F:18-28/Domain: activation peptide #status predicted <ACT>  
F:29-270/Product: (or 31-270) pancreatic elastase IIIB #status predicted <ACT>  
F:29-263/Domain: trypsin homology <TRY>  
F:73,123,217/Active site: His, Asp, Ser #status predicted  
F:114/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F:153/Binding site: carbohydrate (Asn) (covalent) #status absent

Query Match 38.5%; Score 481.5; DB 2; Length 270;  
Best Local Similarity 42.7%; Pred. No. 1.8e-34;  
Matches 103; Conservative 39; Mismatches 86; Indels 13; Gaps 9;

QY 1 IVNGENAVPGSWPQVSLQ-DNTG--FHFCGGSLISPNVWVTAACHQVTPGRHFVVLGEY 57  
Db 29 VVNGEDAVPYSPWQVSLQYKESGSLYHTCGGSLIADPWVVTAGHCILSSRTYQVVLGEY 88

QY 58 DRSSNAEPVQVLSI--ARAITHPNWNNNTM--NNDTLTKLASPARYTAQVSPVCLASTN 113  
Db 89 DRAVKEGPPQVPIPSNGDLFVHPLWNRSCVACGNDIALIKLSRQAGDVLASLPAG 148  
QY 114 EALPSGLICVTGWRISGVGNVTPARLQVVLPLVTYNQCQ--YWGARIIDAMICAGG 171  
Db 149 DILPNETPCYITGWRGLYTNGLP-PDKLOEALLPVVDYEHCSRNNWWSVSKTKMVCAGG 207  
QY 172 SCAGSCGDSGGPLVC-QKGNFWLIGIVSW-GTRKNCNI-QAPAMYTRYSKFSTWVNOVM 228  
Db 208 DIRSGCNGDSGGPLNCPTEDGGQVHGVTFSVSAFCNTRRKPVTFTVTRYSAFIDWIEETI 267  
QY 229 A 229  
Db 268 A 268

RESULT 15  
A25528  
pancreatic elastase II (EC 3.4.21.71) precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 22-Jun-1999  
C:Accession: A25528  
R:Stevenson, B.J.; Hagenbuechle, O.; Wellauer, P.K.  
Nucleic Acids Res. 14, 8307-8330, 1986  
A:Title: Sequence organisation and transcriptional regulation of the mouse elastase I  
A:Reference number: A93646; MUID:87066713; PMID:3641189  
A:Accession: A25528  
A:Molecule type: mRNA  
A:Residues: 1-271 <STE>  
A:Cross-references: GB:X04573; NID:950825; PIDN:CAA28242.1; PID:g50826  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; serine proteinase  
F:1-30/Domain: signal sequence #status predicted <SIG>  
F:31-271/Product: pancreatic elastase II #status predicted <MAT>  
F:31-264/Domain: trypsin homology <TRY>  
F:75,123,218/Active site: His, Asp, Ser #status predicted

Query Match 37.9%; Score 475; DB 2; Length 271;  
Best Local Similarity 40.5%; Pred. No. 6.5e-34;  
Matches 98; Conservative 39; Mismatches 93; Indels 12; Gaps 8;

QY 1 IVNGENAVPGSWPQVSLQ---DNTGHHFCGGSLISPNVWVTAACHQVTPGRHFVVLGEY 57  
Db 31 VVGGQEAETPTWQVSLQVLSLSSGRWRHNCGGSLVANNWVLTAAHCLSNYQTYRVLIGAH 90  
QY 58 DRSSNAEPVQVLSIARAITHPNWNNNTM--DLTKLASPARYTAQVSPVCLASTNEA 115  
Db 91 SLSPNPGAGSAVQVSKLVVHQRWNSONGYDIALIKLASPVTLTKNIQTACLPGAGTI 150  
QY 116 LPSSGLTCVTGWRISGVGNVTPARLQVVLPLVTYNQC--ROYWGARITDMICAGGSG 173  
Db 151 LPRNYCYVTGWLQLTNGN-SPDTRLQGRLLVVDYATCSSASWWSVSKSSMVCAGGDG 209  
QY 174 -ASSCGDSGGPLVCOKGN-TWVLIGIVSWGTK-NCNI-QAPAMYTRYSKFSTWVNOVMA 229  
Db 210 VTSSCNGDSGGPLNCRASNGQVHGIVSFGSLGCGNPKRPSVFTVRSNYIDWINSVMA 269  
QY 230 YN 231  
Db 270 RN 271

Search completed: December 20, 2002, 15:19:50  
Job time : 15.6087 secs



Result No.	Score	Query			DB	ID	Description
		Match	Length	Length			
1	1252	100.0	264	21	AAB11711	Mouse seri	
2	1112	88.8	264	21	AAB11710	Human seri	
3	756	60.4	192	21	AAB54077	Human panc	
4	755.5	60.3	263	23	AAU82738	Amino acid	
5	750.5	59.9	231	22	AAB98504	Human chym	
6	727.5	58.1	245	21	AAV99596	Bovine chym	
7	503	40.2	269	7	AAFG1724	Porcine ela	
8	503	40.2	269	8	AAFG0758	Pig pancrea	
9	500.5	40.0	253	13	AAAR2961	Porcine pa	
10	486	38.8	269	7	AAFG0062	Sequence of	

PT Serine proteases BSSP5, useful in detecting homologs, mutants and

PT polymorphic variants as markers for diagnosis of e.g. Alzheimer's  
 PT disease, epilepsy, cancer and inflammation, using blood, urine,  
 PT pancreas or other tissues  
 XX  
 XX Claim 3; Page 55-56; 70pp; Japanese.  
 XX  
 CC The invention relates to novel serine proteases designated BSSP5  
 CC (AAB11710-B11711), and to nucleic acids encoding them (AAA61733-A61734).  
 CC The invention also relates to vectors and transformants comprising BSSP5  
 CC nucleic acids; transgenic animals in which the expression level of BSSP5  
 CC can be varied; and an mBSSP5 knockout mouse. The invention additionally  
 CC encompasses anti-BSSP5 antibodies and methods of production of such  
 CC antibodies, methods of BSSP5 detection using the antibodies, and the  
 CC use of BSSP5 proteins or fragments as diagnostic markers for certain  
 CC medical conditions, e.g., pancreatitis. A method for detecting  
 CC pancreatitis comprising measuring BSSP5 concentration in the blood or  
 CC urine, and a pancreatitis diagnostic agent containing an anti-BSSP5  
 CC antibody is also disclosed. Nucleotides encoding BSSP5 were initially  
 CC isolated in a human brain cDNA library using degenerate PCR primers  
 CC (AAA61744-A61745) based on conserved regions of serine proteases. The  
 CC BSSP5 serine proteases and nucleotides encoding them are useful in  
 CC detecting homologues, mutants and polymorphic variants in biological  
 CC samples (e.g., blood, urine, brain, prostate gland, placenta, testis,  
 CC pancreas and spleen) as diagnostic markers for conditions such as  
 CC Alzheimer's disease, epilepsy, cancer, inflammation, infertility,  
 CC pancreatitis and prostatic hypertrophy. Sequence AAB11710 represents  
 CC human BSSP5 (hbSSP5), and sequence AAB11711 represents murine BSSP5  
 CC (mBSSP5).  
 XX  
 XX Sequence 264 AA;  
 XX  
 Query Match 100.0%; Score 1252; DB 21; Length 264;  
 Best Local Similarity 100.0%; Pred. No. 3e-102;  
 Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 IVNGENAVPGSWPQVSLQDNTGFHFCGGLISPNVWVTAACHQVTPGRHFVVLGEYDRS 60  
 DB 34 IVNGENAVPGSWPQVSLQDNTGFHFCGGLISPNVWVTAACHQVTPGRHFVVLGEYDRS 93  
 QY 61 SNAEPQVLSIARAIHPNNANTMNDLTLKLSAPRYTAQVSPVCLASTNEALPSGL 120  
 DB 94 SNAEPQVLSIARAIHPNNANTMNDLTLKLSAPRYTAQVSPVCLASTNEALPSGL 153  
 QY 121 TCVTGWRISGVGNVTPARLQVPLVTVNOCROYWGARITDAMICAGGSSCQGD 180  
 DB 154 TCVTGWRISGVGNVTPARLQVPLVTVNOCROYWGARITDAMICAGGSSCQGD 213  
 QY 181 SGGPLVCQKNTWVLIGIVSWGTKNCNIQAPAMYTRVSKFSTWVNOVMAYN 231  
 DB 214 SGGPLVCQKNTWVLIGIVSWGTKNCNIQAPAMYTRVSKFSTWVNOVMAYN 264  
 RESULT 2  
 AAB11710  
 ID AAB11710 standard; Protein; 264 AA.  
 XX  
 XX AAB11710;  
 XX  
 XX 23-OCT-2000 (first entry)  
 XX  
 DE Human serine protease BSSP5 (hbSSP5) SEQ ID NO:2.  
 KW BSSP5; serine protease; human; hbSSP5; mouse; mBSSP5; brain;  
 KW diagnostic marker; antibody; transgenic animal; Alzheimer's disease;  
 KW epilepsy; cancer; inflammation; infertility; pancreatitis;  
 KW prostatic hypertrophy.  
 XX  
 XX Homo sapiens.  
 XX  
 XX WO200031243-A1.  
 XX  
 XX 02-JUN-2000.  
 XX

PF 19-NOV-1999; 99WO-JP06473.  
 XX  
 PR 20-NOV-1998; 98JP-0347806.  
 XX  
 XX (FUSO ) FUSO PHARM IND LTD.  
 XX  
 XX Uemura H, Okui A, Kominami K, Yamaguchi N, Mitsui S;  
 PI  
 XX WPI: 2000-400058/34.  
 DR N-PSDB; AAA61733.  
 XX  
 PT Serine proteases BSSP5, useful in detecting homologs, mutants and  
 PT polymorphic variants as markers for diagnosis of e.g. Alzheimer's  
 PT disease, epilepsy, cancer and inflammation, using blood, urine,  
 PT pancreas or other tissues  
 XX  
 XX Claim 1; Page 51-52; 70pp; Japanese.  
 PS  
 CC The invention relates to novel serine proteases designated BSSP5  
 CC (AAB11710-B11711), and to nucleic acids encoding them (AAA61733-A61734).  
 CC The invention also relates to vectors and transformants comprising BSSP5  
 CC nucleic acids; transgenic animals in which the expression level of BSSP5  
 CC can be varied; and an mBSSP5 knockout mouse. The invention additionally  
 CC encompasses anti-BSSP5 antibodies and methods of production of such  
 CC antibodies, methods of BSSP5 detection using the antibodies, and the  
 CC use of BSSP5 proteins or fragments as diagnostic markers for certain  
 CC medical conditions, e.g., pancreatitis. A method for detecting  
 CC pancreatitis comprising measuring BSSP5 concentration in the blood or  
 CC urine, and a pancreatitis diagnostic agent containing an anti-BSSP5  
 CC antibody is also disclosed. Nucleotides encoding BSSP5 were initially  
 CC isolated in a human brain cDNA library using degenerate PCR primers  
 CC (AAA61744-A61745) based on conserved regions of serine proteases. The  
 CC BSSP5 serine proteases and nucleotides encoding them are useful in  
 CC detecting homologues, mutants and polymorphic variants in biological  
 CC samples (e.g., blood, urine, brain, prostate gland, placenta, testis,  
 CC pancreas and spleen) as diagnostic markers for conditions such as  
 CC Alzheimer's disease, epilepsy, cancer, inflammation, infertility,  
 CC pancreatitis and prostatic hypertrophy. Sequence AAB11710 represents  
 CC human BSSP5 (hbSSP5), and sequence AAB11711 represents murine BSSP5  
 CC (mBSSP5).  
 XX  
 XX Sequence 264 AA;  
 XX  
 Query Match 88.8%; Score 1112; DB 21; Length 264;  
 Best Local Similarity 85.7%; Pred. No. 6.3e-90;  
 Matches 198; Conservative 22; Mismatches 11; Indels 0; Gaps 0;  
 QY 1 IVNGENAVPGSWPQVSLQDNTGFHFCGGLISPNVWVTAACHQVTPGRHFVVLGEYDRS 60  
 DB 34 IVNGENAVPGSWPQVSLQDNTGFHFCGGLISQSWVWVTAACHCNVSPGRHFVVLGEYDRS 93  
 QY 61 SNAEPQVLSIARAIHPNNANTMNDLTLKLSAPRYTAQVSPVCLASTNEALPSGL 120  
 DB 94 SNAEPQVLSIARAIHPNNANTMNDLTLKLSAPRYTAQVSPVCLASTNEALPSGL 153  
 QY 121 TCVTGWRISGVGNVTPARLQVPLVTVNOCROYWGARITDAMICAGGSSCQGD 180  
 DB 154 TCVTGWRISGVGNVTPARLQVPLVTVNOCROYWGARITDAMICAGGSSCQGD 213  
 QY 181 SGGPLVCQKNTWVLIGIVSWGTKNCNIQAPAMYTRVSKFSTWVNOVMAYN 231  
 DB 214 SGGPLVCQKNTWVLIGIVSWGTKNCNVRAPAVYTRVSKFSTWVNOVMAYN 264  
 RESULT 3  
 AAB54077  
 ID AAB54077 standard; Protein; 192 AA.  
 XX  
 XX AAB54077;  
 XX  
 XX 09-MAR-2001 (first entry)  
 XX  
 DE Human pancreatic cancer antigen protein sequence SEQ ID NO:529.

XX Human; pancreas; pancreatic cancer; pancreatic cancer antigen;  
KW detection; diagnosis; identification; cytostatic; neuroprotective;  
KW nontropic; immunomodulatory; relaxant; contraceptive; gynaecological;  
KW antiinflammatory; cardiant; gene therapy; chromosome mapping;  
KW linkage analysis; tissue identification; tissue typing; forensic;  
KW neural; immune system; muscular; reproductive; gastrointestinal;  
KW pulmonary; cardiovascular; renal; proliferative.  
XX  
OS Homo sapiens.  
XX  
XX  
PN WO200055320-A1.  
XX  
XX 21-SEP-2000.  
XX  
XX 08-MAR-2000; 2000WO-US05989.  
XX  
XX 12-MAR-1999; 99US-0124270.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Ruben SM;  
XX  
XX WPI; 2000-579444/54.  
DR N-PSDB; AAC98842.  
XX  
XX New nucleic acid that is a pancreatic cancer antigen for preventing,  
PT treating, or ameliorating a medical condition, particular pancreatic  
PT cancer, or for use in assays for diagnosing a pathological condition -  
XX  
XX Claim 11; Page 966; 1379pp; English.  
XX  
XX AAC98773 to AAC9231 encode the human pancreatic cancer associated  
CC proteins, called pancreatic cancer antigens, given in AAB54008 to  
CC AAB54466. The human pancreatic cancer antigens have cytostatic,  
CC neuroprotective, nontropic, immunomodulatory, relaxant, contraceptive,  
CC gynaecological, cardiant and antiinflammatory activities, and can be used  
CC in gene therapy. The polynucleotide and proteins can be used for  
CC preventing, treating, or ameliorating a medical condition or in assays  
CC for diagnosing a pathological condition or a susceptibility to one in a  
CC subject. Binding partners to the proteins and the activity of the  
CC proteins can be identified. The pancreatic cancer antigens can be used to  
CC detect, treat or prevent pancreatic disorders, especially cancer.  
CC Agonists and antagonists to the antigens can be screened for. The  
CC pancreatic cancer antigen polynucleotides can be used to design nucleic  
CC acid hybridisation probes that can be used in chromosome mapping, linkage  
CC analysis, tissue identification and/or typing and a variety of forensic  
CC and diagnostic methods. The proteins can be used to generate antibodies  
CC which are used to purify, detect and target the polypeptides, including  
CC both in vivo and in vitro diagnostic and therapeutic methods. The  
CC proteins can be used to treat or prevent neural, immune system, renal or  
CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or  
CC proliferative disorders. AAC9232 to AAC9240 and AAB54467 represent  
CC sequences used in the exemplification of the present invention.  
XX  
XX Sequence 192 AA;  
Query Match 60.4%; Score 756; DB 21; Length 192;  
Best Local Similarity 84.4%; Pred. No. 9.2e-59;  
Matches 135; Conservative 16; Mismatches 9; Indels 0; Gaps 0;  
QY 1 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNVVTAAHCQVTPGRHFVVLGEYDRS 60  
Db 30 IVNGENAVLGSWPQVSLQDSSGFHFCGSLISQSNVVTAAHCNVSPGRHFVVLGEYDRS 89  
QY 61 SNAEPQVLSIARATHPHNNANTMNDTLTKLASPARYTAAQVSPVCLASTNEALPSGL 120  
Db 90 SNAEPQLVLSVRAITHPSWNTNDVTLTKLASPAQYTTTRISPVCLASTNEALTEGL 149  
QY 121 TCVTGWRISGVNVTPLAQVVLPLVTVNCQRYWGA 160  
Db 150 TCVTGWRISGVNVTPLAQVVLPLVTVNCQRYWGS 189

## RESULT 4

AAU82738  
XX AAU82738 standard; Protein; 263 AA.  
AC AAU82738;  
XX  
XX 23-APR-2002 (first entry)  
XX  
XX Amino acid sequence of novel human protease #37.  
XX  
XX Human; protease; cancer; immune-related disorder; cardiovascular disease;  
KW neuronal-associated disease; metabolic disorder; inflammatory disorder;  
KW nervous system disorder; sexual dysfunction; pain; mood disorder;  
KW hypertension; psychotic disorder; neurological disorder; dyskinesia;  
KW viral infection; human immunodeficiency virus; HIV; non-viral infection;  
KW ocular disease; cytostatic; enzyme.  
XX  
XX Homo sapiens.  
XX  
XX WO200200860-A2.  
XX  
XX 03-JAN-2002.  
XX  
XX 26-JUN-2001; 2001WO-US20171.  
XX  
XX 26-JUN-2000; 2000US-214047P.  
XX  
XX (SUGE-) SUGEN INC.  
XX  
XX Plowman G, Whyte D, Sudarsanam S, Manning G, Caenepeel S;  
PI Charyczak G;  
XX  
XX WPI; 2002-139913/18.  
DR N-PSDB; ABK31780.  
XX  
XX Nucleic acids encoding novel human proteases, useful for useful for  
PT treating diseases and disorders such as cancers, immune-related  
PT diseases and disorders, cardiovascular disease (e.g. restenosis) and  
PT inflammatory disorders -  
XX  
XX Claim 6; Fig 2M; 313pp; English.  
XX  
XX The present invention relates to the isolation of novel human  
CC proteases, and the nucleic acids encoding them. The sequences of  
CC the invention are useful for treating diseases and disorders such as  
CC cancers (e.g. breast, colon, lung), immune-related diseases and disorders  
CC (e.g. inflammatory diseases and asthma), cardiovascular diseases  
CC (e.g. restenosis and coronary thrombosis), brain or neuronal-associated  
CC diseases, metabolic disorders (e.g. diabetes, obesity), inflammatory  
CC disorders (e.g. rheumatoid arthritis and psoriasis), central or  
CC peripheral nervous system diseases, migraines, pain, sexual dysfunction,  
CC mood disorders, attention disorders, cognition disorders, hypotension,  
CC hypertension, psychotic disorders, neurological disorders  
CC (e.g. Alzheimer's disease, Parkinson's disease) and dyskinesias.  
CC The nucleic acids and polypeptides are also useful for treating viral  
CC infections caused by human immunodeficiency virus (HIV), and non-viral  
CC infections such as ocular disease (e.g. glaucoma) and macular  
CC degeneration. AAU82702-AAU82760 represent the novel human proteases of  
CC the invention.  
XX  
XX Sequence 263 AA;  
Query Match 60.3%; Score 755.5; DB 23; Length 263;  
Best Local Similarity 58.4%; Pred. No. 1.5e-58;  
Matches 135; Conservative 35; Mismatches 60; Indels 1; Gaps 1;  
QY 1 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNVVTAAHCQVTPGRHFVVLGEYDRS 60  
Db 34 IVNGEDAVPGSWPQVSLQDKTGFHFCGSLISEDVVTAAHCGVRTS-DVVVAGEFDG 92  
QY 61 SNAEPQVLSIARATHPHNNANTMNDTLTKLASPARYTAAQVSPVCLASTNEALPSGL 120  
Db 150 SNAEPQVLSIARATHPHNNANTMNDTLTKLASPARYTAAQVSPVCLASTNEALPSGL 120

Db 93 SDEENIOVLKIAKFKNPKFSILTVNNDITLLKLTATPAREFSQTVSAVCLPSADDDFFAGT 152  
QY 121 TCVTTGWRISGVGNVTPARLQOVVPLVTVNOCROYWGARITDAMICAGSGASSCGQD 180  
Db 153 LCATTGKGKYNANKTPDKLQQAALPLLSNAECKSKWGRITDVMICAGSGVSSCMGD 212  
QY 181 SGGPLVCQKGNWTLIGIVSGTKNCNIQAPAMTRYSKFSTWINOVMAYN 231  
Db 213 SGGPLVCQKGCAMWTLVGIWSGSRCTSTTPAVYARVTKLIPWQKILAA 263

RESULT 5  
AAB98504  
ID AAB98504 standard; Protein; 231 AA.  
XX AC AAB98504;  
DT 03-AUG-2001 (first entry)  
DE Human chymotrypsin serine protease catalytic domain.  
KW Human; TADG-15; cytostatic; vaccine; ovarian tumour; cancer;  
KW tumour antigen-derived gene 15; serine protease; chymotrypsin.  
XX Homo sapiens.  
XX WO200129056-A1.  
XX 26-APR-2001.  
XX 20-OCT-2000; 2000WO-US29095.  
XX 20-OCT-1999; 99US-0421213.  
XX (UYAR-) UNIV ARKANSAS.  
XX O'Brien TJ, Tanimoto H;  
XX WPI; 2001-381031/40.  
XX Novel extracellular serine protease, termed tumor antigen-derived gene  
PT 15 protein overexpressed in carcinomas and DNA encoding it, for  
PT diagnosis, treatment, prevention of cancer, particularly breast,  
PT ovarian cancer -  
XX Example 10; Fig 1; 130pp; English.  
XX The present invention relates to human tumour antigen-derived gene 15  
CC (TADG-15) protein and coding sequence (see AAB23601 and AAB98500).  
CC TADG-15 is an extracellular serine protease. It was found that TADG-15 is  
CC over-expressed in ovarian tumours. TADG-15 protein or its fragments of  
CC 9-20 residues that lack TADG-15 protease activity are useful for  
CC vaccinating an individual against TADG-15, having, suspected of having or  
CC at risk of getting cancer. Furthermore, the TADG-15 gene can be used as a  
CC diagnostic or therapeutic target in cancer. The present sequence was used  
CC in a sequence homology alignment with the catalytic domain of TADG-15.  
XX  
SQ Sequence 231 AA;

Query Match 59.9%; Score 750.5; DB 22; Length 231;  
Best Local Similarity 58.0%; Pred. No. 3.5e-58;  
Matches 134; Conservative 35; Mismatches 61; Indels 1; Gaps 1;

QY 1 IVNGENAVPGSPWQVSLQDNTGFHFCGSLISPNWVYTAHCOVTPGRHFWVLGEYDRS 60  
Db 2 IVNGEDAVPGSPWQVSLQDNTGFHFCGSLISDNWVYTAHCOVTPGRHFWVLGEYDRS 60  
QY 61 SNAEPVOVLSTARAITHPNNANMNDLTLKLSAPRYTAQVSPVCLASTNEALPSGL 120  
Db 61 SDEENIOVLKIAKFKNPKFSILTVNNDITLLKLTATPAREFSQTVSAVCLPSADDDFFAGT 120  
QY 121 TCVTTGWRISGVGNVTPARLQOVVPLVTVNOCROYWGARITDAMICAGSGASSCGQD 180  
Db 121 SGGPLVCQKGCAMWTLVGIWSGSRCTSTTPAVYARVTKLIPWQKILAA 263

Db 121 LCATTGKGKYNANKTPDKLQQAALPLLSNAECKSKWGRITDVMICAGSGVSSCMGD 180  
QY 181 SGGPLVCQKGNWTLIGIVSGTKNCNIQAPAMTRYSKFSTWINOVMAYN 231  
Db 181 SGGPLVCQKGCAMWTLVGIWSGSDTCTSPGTVARVTKLIPWQKILAA 231

RESULT 6  
AAY99596  
ID AAY99596 standard; protein; 245 AA.  
XX AC AAY99596;  
DT 13-SEP-2000 (first entry)  
DE Bovine chymotrypsinogen A.  
KW Bovine; plasminogen activator; cardiant; thrombolytic;  
KW heart attack; stroke; blood clotting disorder.  
XX Bos taurus.  
XX WO200032759-A1.  
XX 08-JUN-2000.  
XX 06-MAY-1999; 99WO-US09991.  
XX 02-DEC-1998; 98US-0110588.  
XX (OKLA-) OKLAHOMA MEDICAL RES FOUND.  
XX Lin X, Zhang XC, Tang JUN;  
XX WPI; 2000-422975/36.  
XX Polypeptide with plasminogen activator activity useful as thrombolytic  
PT agent for treating blood clot disorders e.g. heart attack, comprises 10  
PT amino acid peptide fragment for recognition or activation of  
PT plasminogen -  
XX Disclosure; Page 35-36; 41pp; English.  
XX The present sequence is bovine chymotrypsinogen. It was included in  
CC a review of sequence homologies of several plasminogen activators.  
CC Plasminogen is the principal serine protease zymogen in the  
CC extracellular fluids of vertebrates. Its active form, plasmin, is  
CC implicated in pericellular proteolysis associated with a wide range of  
CC physiological and pathological processes. Plasminogen expression is  
CC regulated by plasminogen activators which hydrolyse a peptide bond in  
CC plasminogen to convert it to plasmin or form tight binding complexes  
CC with plasminogen to spontaneously convert it to plasmin. The sequence  
CC homology analysis has identified a six amino acid peptide involved in  
CC plasminogen activation. This peptide is particularly useful when inserted  
CC between amino acid residues 644 and 645 of full length human  
CC plasminogen. Novel plasminogen activators have been made based upon the  
CC plasminogen activation/recognition site of plasminogen binding proteins.  
CC The polypeptides are useful in preparing thrombolytic agents for treating  
CC blood clotting disorders such as heart attack.  
XX  
SQ Sequence 245 AA;

Query Match 58.1%; Score 727.5; DB 21; Length 245;  
Best Local Similarity 56.7%; Pred. No. 4e-56;  
Matches 131; Conservative 32; Mismatches 67; Indels 1; Gaps 1;

QY 1 IVGENAVPGSPWQVSLQDNTGFHFCGSLISPNWVYTAHCOVTPGRHFWVLGEYDRS 60  
Db 16 IVNGEAVPGSPWQVSLQDNTGFHFCGSLISDNWVYTAHCOVTPGRHFWVLGEYDRS 74  
QY 61 SNAEPVOVLSTARAITHPNNANMNDLTLKLSAPRYTAQVSPVCLASTNEALPSGL 120  
Db 75 SSSEKIQLKIAKFKNPKFSILTVNNDITLLKLTATPAREFSQTVSAVCLPSADDDFAAGT 134

Query Match	Best Local Similarity	40.2%	Score 503
Matches 107;	Conservative 41;	Mismatch	
Qy	1	IVNGENAVPGSMPQWVSLQ-DNTG--FHFCGGS	
Db	29	VGQGDARNSMPQWVSLQYDSSGQWRHCTGGT	
Qy	58	DRSSNAEPQVQLSIARAIHPNNANTMN--ND	
Db	89	SLSTNEPGSLAVKVSLLVHQDANSNQLSNGND	
Qy	116	LPSSLTCVTGTRIGSGVGNVTPIRQQVVLPL	
Db	149	LPNNYYCYVTGWRGLQTNG-ASPDLIQQQQLLV	
Qy	174	-ASSCCGGDGGPLVLCQKGN-TWVLIGIVSWGPK	
Db	208	IISCCNGDGGPLNCGANGQWGHVIGIVSFGSS	
Qy	230	YN 231	
Db	268	NN 269	
RESULT 9			
AAR29621			
ID AAR29621	standard; Protein; 253 AA.		

```

XX AAR29621;
AC
XX
XX 30-APR-1993 (first entry)
DT
XX
XX Porcine pancreatic elastase III.
DE
XX
XX Pig; swine; bile acid; gall bladder; bile acid secretion promoters;
KW liver function improvers; ss.
XX
XX Sus scrofa domestica.
OS
XX
XX JP04325090-A.
PN
XX
XX 13-NOV-1992.
PD
XX
XX 23-APR-1991; 91JP-0092069.
PF
XX
XX 23-APR-1991; 91JP-0092069.
PR
XX
XX (SANY ) SANKYO CO LTD.
PA
XX
XX WPI; 1992-428826/52.
DR
XX
XX N-PSDB; AAQ31724.
DR
XX
XX Pig pancreas elastase protein - used in bile acid secretion
PT promoters and liver function improvers
PT
XX
XX Claim 1; Page 8; 10pp; Japanese.
PS
XX
XX This sequence represents porcine pancreatic elastase III. It was
CC expressed in E. coli YA21 using the expression vector pELE001.
CC It may be used in bile acid secretion promoters and liver function
CC improvers.
XX
XX Sequence 253 AA;
Query Match 40.0%; Score 500.5; DB 13; Length 253;
Best Local Similarity 43.2%; Pred. No. 3.9e-36;
Matches 104; Conservative 42; Mismatches 82; Indels 13; Gaps 9;

QY 1 IVNGENAVPGSPWQVSLQ--DNTGF--HFCGGSLISPNWVYTAARHCQVTPGRHFVVLGEY 57
DB 12 VNGEDAVPYSPWQVSLQEKNGVFQHTCGGSLIAPDWLTAGHCISSSLTYQVVLGEY 71
QY 58 DRSSNAEPVQVLSI--ARAITHPNWNTMN--NDLTLLKLPARYTAQVSPVCLASTN 113
DB 72 DRSENEGFEQVIPINAGDLFVHPRWNSVCVSCGNDIALVKLSAQLGDKVOLACLPPAG 131
QY 114 EALPSGLTCVTTGWGRISGVGNVTPARLQOVPLVTVNOCRO--YWGARITDAMICAGG 171
DB 132 DILPNDTPCISGWRGLYTNGLP--PKLQALLPVVDYQHCSDWWDWGSTVKTWTCAGG 190
QY 172 SGASSCGDGGPLVCQKGN--TWVLIGIVSW--GTKNCN--IQAPAMYTRVSKFSTWINQVM 228
DB 191 DIRSGCNGDGGPLNCPAADSGMQVHGVTFSVSAYGCTNLKPKPTVTRTSAFIDWIEEII 250
QY 229 A 229
DB 251 A 251

RESULT 10
AAP60062
ID AAP60062 standard; Protein; 269 AA.
XX
XX AAP60062;
AC
XX
XX 06-SEP-1991 (first entry)
DT
XX
XX Sequence of human pancreatic elastase IIA encoded on pH2E2.
DE
XX
XX Enzyme; serum lipoprotein metabolism.
KW
XX

```

```

XX Homo sapiens.
OS
XX
XX EP198645-A.
PN
XX
XX 22-OCT-1986.
PD
XX
XX 07-APR-1986; 86EP-0302557.
PF
XX
XX 23-OCT-1985; 85JP-0236686.
PR
XX
XX 05-APR-1985; 85JP-0072308.
PR
XX
XX 27-APR-1985; 85JP-0091986.
PR
XX
XX 26-JUL-1985; 85JP-0163964.
PR
XX
XX 02-DEC-1985; 85JP-0271128.
PA
XX
XX (SANY ) SANKYO KK.
XX
XX Takiguchi Y, Tani T, Kawashima I, Erukawa H, Ohmine T;
PI Ohsumi J;
XX
XX WPI; 1986-280300/43.
DR
XX
XX N-PSDB; AAN60076.
DR
XX
XX New genetically engineered human pancreatic elastase - obtd.
PT using hosts modified DNA coding for enzyme
PT
XX
XX Example; Page 14-15; 45pp; English.
PS
XX
XX The genetically engineered prod. can eliminate the dependency on
CC human pancreas supplies for the elastase, and avoids antibody
CC formation and possibility of anaphylaxis using porcine elastase.
XX
XX Sequence 269 AA;
Query Match 38.8%; Score 486; DB 7; Length 269;
Best Local Similarity 43.0%; Pred. No. 8e-35;
Matches 104; Conservative 41; Mismatches 85; Indels 12; Gaps 8;

QY 1 IVNGENAVPGSPWQVSLQDNTG--FHCGGSLISPNWVYTAARHCQVTPGRHFVVLGEY 57
DB 29 VVGEEARPNSPWQVSLQYSSNGKWYHTCGGSLIANSWVLTAAHCISSSRTYVGLGRH 88
QY 58 DRSSNAEPVQVLSIARAITHPNWNTMN--NDLTLLKLPARYTAQVSPVCLASTNEA 115
DB 89 NLVVAESGSLAVSVSKIVVHKDWNQISKGNDAIKLANPVSLTDKIQLACLPPAGTI 148
QY 116 LPSGLTCVTTGWGRISGVGNVTPARLQOVPLVTVNOCRO--YWGARITDAMICAGGSG 173
DB 149 LPNYPCTVTGWRLQTNQAV--PDVLQOGRLLVVDYATCSSSAMWSSSVKTSMICAGDG 207
QY 174 A--SSCGDGGPLVCQKGN--TWVLIGIVSWGK--NCN--IQAPAMYTRVSKFSTWINQVM 229
DB 208 VISCCNGDGGPLNCPAADSGMQVHGIVSPGRLGCNYYHKPSVFTVRSNYIDWINSVIA 267
QY 230 YN 231
DB 268 NN 269

RESULT 11
AAP60061
ID AAP60061 standard; Protein; 242 AA.
XX
XX AAP60061;
AC
XX
XX 06-SEP-1991 (first entry)
DT
XX
XX Sequence of human pancreatic elastase IIIB.
DE
XX
XX Enzyme; serum lipoprotein metabolism.
KW
XX
XX Homo sapiens.
OS
XX

```

```

PN EPI98645-A.
XX PD 22-OCT-1986.
XX PF 07-APR-1986; 86EP-0302557.
XX PR 23-OCT-1985; 85JP-0236686.
XX PR 05-APR-1985; 85JP-0072308.
XX PR 27-APR-1985; 85JP-0091986.
XX PR 26-JUL-1985; 85JP-0163964.
XX PR 02-DEC-1985; 85JP-0271128.
XX PA (SANY ) SANKYO KK.
XX Takiguchi Y, Tani T, Kawashima I, Eurukawa H, Ohmine T;
PI Ohsumi J;
XX WPI; 1986-280300/43.
XX DR N-PSDB; AAN60075.
XX New genetically engineered human pancreatic elastase - obt'd.
XX PT using hosts modified DNA coding for enzyme
XX PS Claim 2; Page 8; 45pp; English.
XX The genetically engineered prod. can eliminate the dependency on
CC human pancreas supplies for the elastase, and avoids antibody
CC formation and possibility of anaphylaxis using porcine elastase.
XX SQ Sequence 242 AA;
Query Match 38.5%; Score 481.5; DB 7; Length 242;
Best Local Similarity 42.7%; Pred. No. 1.7e-34;
Matches 103; Conservative 39; Mismatches 86; Indels 13; Gaps 9;
Qy 1 IVNGENAVPGSWPQVLSIQ-DNTG--FHFCGSLISPNVYTAHCQVTPGRHFVVLGEY 57
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1 VVNGEDAVPYSWPQVLSIQEYKSGSYHTCGGSLIAPDWMVYTAGHCISSSRTYQVVLGEY 60
Qy 58 DRSSNAEPVQVLSI--ARAITHPNWNTM--NDLTLLKLASPARYTAQVSPVCLASTN 113
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
61 DRVKEGPEQVPIPSGDLFVHPLNRSVACGNDIALIKLSRQGLDQVQLASLPAG 120
Qy 114 EALPSGLTCVTGMRISGVGNVTPARLQVVLPLVTYNOCRQ--YNGARITDMICAGG 171
Db :||:| | ||||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
121 DILNETECYITGWRGLVTNGPL-PDKLQEALLPVVDYEHCSRNNWCGSSVKKTWVCAGG 179
Qy 172 SGASSCGDGGPLVC-OKGNTWVLIGIVSW-GTKNCNI-QAPAMYTRYSKFSTWINQVM 228
Db | | ||||| | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
180 DIRSGCNGDGGPLNCPTEDGGWQVHGVTSFVSFAGCNTRRKPTVTRVSADFIDWIBETI 239
Qy 229 A 229
Db 240 A 240
RESULT 12
AAP60058
ID AAP60058 standard; Protein; 252 AA.
XX AC AAP60058;
XX 06-SEP-1991 (first entry)
XX DE Sequence of human pancreatic elastase IIA.
XX KW Enzyme: serum lipoprotein metabolism.
XX OS Homo sapiens.
XX PN EPI98645-A.
XX PD 22-OCT-1986.

```

```

XX 07-APR-1986; 86EP-0302557.
XX 23-OCT-1985; 85JP-0236686.
XX 05-APR-1985; 85JP-0072308.
XX 27-APR-1985; 85JP-0091986.
XX 26-JUL-1985; 85JP-0163964.
XX 02-DEC-1985; 85JP-0271128.
XX PA (SANY ) SANKYO KK.
XX Takiguchi Y, Tani T, Kawashima I, Eurukawa H, Ohmine T;
PI Ohsumi J;
XX WPI; 1986-280300/43.
XX DR N-PSDB; AAN60072.
XX New genetically engineered human pancreatic elastase - obt'd.
XX PT using hosts modified DNA coding for enzyme
XX PS Claim 2; Page 3-4; 45pp; English.
XX The genetically engineered prod. can eliminate the dependency on
CC human pancreas supplies for the elastase, and avoids antibody
CC formation and possibility of anaphylaxis using porcine elastase.
XX SQ Sequence 252 AA;
Query Match 37.7%; Score 471.5; DB 7; Length 252;
Best Local Similarity 42.6%; Pred. No. 1.4e-33;
Matches 103; Conservative 41; Mismatches 85; Indels 13; Gaps 9;
Qy 1 IVNGENAVPGSWPQVLSIQDNTG---FHFCGSLISPNVYTAHCQVTPGRHFVVLGEY 57
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
13 VVGGEARPNWPQVLSIQYSSNGKWIYTCGGSLIANGWLTPAHCISSSRTYRVLGRH 72
Qy 58 DRSSNAEPVQVLSIARAITHPNWNTM--NDLTLLKLASPARYTAQVSPVCLASTNEA 115
Db :||:| | ||||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
73 NLYVAESGLAVSVSKIIVVHKDWSNQISKGNIDALLKLANPVSLTKIQLACLPPAGTI 132
Qy 116 LPSGLTCVTGMRISGVGNVTPARLQVVLPLVTYNOCRQ--YNGARITDMICAGGSG 173
Db :||:| | ||||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
133 LPNYPVCYVTGWRGLVTNGAV-PDVLOQRLVAVATCSSAWNGSSVYKTS-ICAGGSG 190
Qy 174 A-SSCGDGGPLVCQKGN-TWVLIGIVSWGTFK-NCN-IOAPAMYTRYSKFSTWINOVMA 229
Db | | ||||| | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
191 VISSCNGDGGPLNCQASDGRQVHGVISFGRSGCNYHKKPSVTRVSNYIDWINSVIA 250
Qy 230 YN 231
Db 251 NN 252
RESULT 13
ABG20513
ID ABG20513 standard; Protein; 279 AA.
XX AC ABG20513;
XX 13-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #20504.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.

```

XX 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX (HYSE-) HYSEQ INC.  
 XX Drmanac RT, Liu C, Tang YT;  
 PI WPI: 2001-639362/73.  
 DR N-PSDB; AAS84700.  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity -  
 XX Claim 20; SEQ ID No 50872; 103pp; English.  
 PS The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABC0010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX Sequence 279 AA;  
 SQ

Query Match 37.6%; Score 471; DB 22; Length 279;  
 Best Local Similarity 41.0%; Pred. No. 1.7e-33;  
 Matches 102; Conservative 39; Mismatches 84; Indels 24; Gaps 9;  
 QY 1 IVNGENAVGSPWPQVSLQDNTG---FHFCGGSLISPNWVTAHC-----QVTPGRH 50  
 Db 37 VVGGEARPNNSWPQVSLQYSSNGKWHYTCGGSLIANSWVLTAAHCISSTRYRVGLGRH 96  
 QY 51 FVYLGEYDRSSNAEPVQLVLSIARAITHPNWNTMN--NDLTLLKLASPARYTAQVSPVC 108  
 Db 97 NLYVAE-----SGSLAMSSVSKIVVHKDWNFNQISKGNIDALLKLANPVS�TDKIQIAL 150  
 QY 109 LASTNEALPSGLTCVTTGGRISGVGNVTPARLQOVVPLVTVNQCRRQ--YWGARITDAM 166  
 Db 151 FPPAGTILPNPYPCVXTGLGESCRTNGAVPDVLOQGRLLVVDYATCSSAWNGSSVKTSM 210  
 QY 167 ICAGGSGA-SSCGDGGPLVCCQKNT--VVLIGIVSGWTK-NCN-IOAPAMYTRVSKFST 222  
 Db 211 ICAGGSGVSSCGDGGPLNCOASDALMQVHGVISFGSLGNCYHKKPSVFTRVSNYID 270  
 QY 223 WINQVMAYN 231  
 Db 271 WINSVIAYN 279  
 RESULT 14  
 AAP61723  
 ID AAP61723 standard; Protein; 269 AA.  
 XX  
 AC AAP61723;  
 XX

DT 23-OCT-1991 (first entry)  
 XX Human elastase II.  
 DE Colibacillus; yeast.  
 XX Homo sapiens.  
 OS JP61192288-A.  
 PN 26-AUG-1986.  
 PD 22-FEB-1985; 85JP-0034049.  
 PF 22-FEB-1985; 85JP-0034049.  
 PR (KIRI ) KIRIN BREWERY KK.  
 PA WPI; 1986-262894/40.  
 DR N-PSDB; AAN60706.  
 XX Biologically producing human elastase ii - using DNA chain with  
 PT biological prodn. capability for human elastase II.  
 PS Claim 1; Fig 1-2; 19pp; Japanese.  
 XX The human elastase product may be efficiently expressed from a  
 CC transformed host such as colibacillus or yeast, yielding the product  
 CC more efficiently than spleen extraction.  
 CC See also J61192289.  
 XX Sequence 269 AA;  
 SQ  
 Query Match 37.5%; Score 469; DB 7; Length 269;  
 Best Local Similarity 42.6%; Pred. No. 2.5e-33;  
 Matches 103; Conservative 40; Mismatches 87; Indels 12; Gaps 8;  
 QY 1 IVNGENAVGSPWPQVSLQDNTG---FHFCGGSLISPNWVTAHCQVTPGRHVVVLGEY 57  
 Db 29 VVGGEARPNNSWPQVSLQYSSNGKWHYTCGGSLIANSWVLTAAHCISSTRYRVGLGRH 88  
 QY 58 DRSSNAEPVQLVLSIARAITHPNWNTMN--NDLTLLKLASPARYTAQVSPVCCLASTNEA 115  
 Db 89 NLYVAESGSLAVSIVVHKDWNFNQISKGNIDALLKLANPVS�TDKIQIALCLPPAGTI 148  
 QY 116 LPSGLTCVTTGGRISGVGNVTPARLQOVVPLVTVNQCRRQ--YWGARITDAMICAGGSG 173  
 Db 149 LPNNYPCVYTGWRGLQTNGAV-PDVLOQGRLLVVDYATCSSAWNGSSVKTSMICAGDGG 207  
 QY 174 A-SSCGDGGPLVCCQKNTWLI-GIVSGWTK-NCN-IOAPAMYTRVSKFSTWINQVMA 229  
 Db 208 VISSCNGDGGPLNCOASDGRCEVHGVISFGSLGNCYHKKPSVFTRVSNIDWINSVIA 267  
 QY 230 YN 231  
 Db 268 NN 269  
 RESULT 15  
 AAW34285  
 ID AAW34285 standard; protein; 791 AA.  
 XX  
 AC AAW34285;  
 XX  
 DT 14-MAY-1998 (first entry)  
 XX Human plasminogen.  
 DE Plasminogen; human; Kringle 5 peptide; anti-angiogenesis agent; cancer;  
 KW metastatic solid tumour; carcinoma; sarcoma; lymphoma; haemangioma;  
 KW psoriasis; arthritis; macular degeneration; diabetic retinopathy;  
 KW autoimmune disease; ocular disease; capillary proliferation; therapy;  
 KW Kringle 5 receptor.







GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:14:08 ; Search time 4.08696 Seconds  
(without alignments)  
1663.721 Million cell updates/sec

Title: US-09-856-319B-4\_COPY\_1\_33

Perfect score: 166

Sequence: 1 MLLSLTSLVLLGSSMGCGVPAITPALSYNQR 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21.\*

1: sp\_archaea.\*

2: sp\_bacteria.\*

3: sp\_fungi.\*

4: sp\_human.\*

5: sp\_invertebrate.\*

6: sp\_mammal.\*

7: sp\_mhc.\*

8: sp\_organelle.\*

9: sp\_phase.\*

10: sp\_plant.\*

11: sp\_rodent.\*

12: sp\_virus.\*

13: sp\_vertebrate.\*

14: sp\_unclassified.\*

15: sp\_rvirus.\*

16: sp\_bacteriap.\*

17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query %	Match	Length	DB ID	Description
1	166	100.0	264	11	Q9D960	Q9d960 mus musculu
2	166	100.0	264	11	Q9D7P8	Q9d7p8 mus musculu
3	166	100.0	264	11	Q9EQZ8	Q9eqz8 rattus norv
4	166	100.0	264	11	Q9ER05	Q9er05 mus musculu
5	71	42.8	263	11	Q9DC86	Q9dc86 mus musculu
6	71	42.8	263	11	Q9D8X8	Q9d8x8 mus musculu
7	71	42.8	263	11	Q9CR35	Q9cr35 mus musculu
8	59.5	35.8	261	13	Q9W7Q4	Q9w7q4 paralichthy
9	59	35.5	648	17	O8THZ0	O8thz0 methanosarc
10	58	34.9	216	4	Q96R18	Q96r18 homo sapien
11	56	33.7	260	13	Q9W7Q3	Q9w7q3 paralichthy
12	55.5	33.4	140	9	Q8SBN3	Q8sbn3 bacterioph
13	55.5	33.4	441	5	Q96666	Q96666 drosophila
14	55.5	33.4	441	5	Q9V3B9	Q9v3b9 drosophila
15	55.5	33.4	835	17	Q97Z55	Q97z55 sulfolobus
16	55	33.1	242	16	Q8ZLR7	Q8zlr7 salmonella

17	55	33.1	242	16	O8Z3G0	O8z3g0 salmonella
18	55	33.1	304	10	Q94AJ4	Q94aj4 arabidopsis
19	55	33.1	465	10	Q9SVB9	Q9sve9 arabidopsis
20	54	32.5	255	10	Q9FW90	Q9fw90 oryza sativ
21	54	32.5	622	16	O8XT78	O8xt78 ralstonia s
22	54	32.5	637	2	Q9RPI8	Q9rpi8 desulfitoba
23	53	31.9	671	10	Q9FEA3	Q9fea3 oryza sativ
24	52	31.3	241	16	O8ZB72	O8zb72 versinia pe
25	52	31.3	292	10	O43752	O43752 citrus sine
26	52	31.3	579	10	O48542	O48542 hordeum vul
27	52	31.3	891	12	Q9J052	Q9j052 simlan herp
28	51.5	31.0	270	13	Q9I039	Q9i039 gadus morhu
29	51.5	31.0	343	2	Q8RSP6	Q8rsp6 rhizobium l
30	51.5	31.0	663	10	O04976	O04976 mangifera 1
31	51	30.7	326	16	O8ZM02	O8zmo2 salmonella
32	51	30.7	533	11	O08569	O08569 cavia porce
33	51	30.7	757	10	O8VXA8	O8vxa8 oryza sativ
34	51	30.7	826	10	O80739	O80739 arabidopsis
35	51	30.7	1186	12	O55767	O55767 chilo iride
36	50.5	30.4	248	6	Q9N0G1	Q9n0g1 equus cabal
37	50.5	30.4	248	6	O95L88	O95l88 equus cabal
38	50.5	30.4	384	16	O8ZF12	O8zf12 yersinia pe
39	50.5	30.4	536	11	O8VHD6	O8vhd6 mus musculu
40	50.5	30.4	538	4	O8WUG5	O8wug5 homo sapien
41	50.5	30.4	1004	16	O9RJT8	O9rjt8 streptomyce
42	50	30.1	94	4	O96F38	O96f38 homo sapien
43	50	30.1	144	5	Q9V316	Q9v316 drosophila
44	50	30.1	169	4	O9H7M4	O9h7m4 homo sapien
45	50	30.1	222	11	Q9JIY6	Q9jiy6 rattus ratt

#### ALIGNMENTS

RESULT 1

Q9D960 ID Q9D960 PRELIMINARY: PRT: 264 AA.  
AC Q9D960;  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE 1810004D15RIK protein.  
OS CTRL OR 1810004D15RIK.  
GS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Schoenwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection."  
RL Nature 409:685-690(2001).  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
TRYPSIN FAMILY.

Q9ER05	PRELIMINARY;	PRT;	264 AA.
ID	Q9ER05		
AC	Q9ER05;		
DT	01-MAR-2001	(TREMBlrel. 16, Created)	
DT	01-MAR-2001	(TREMBlrel. 16, Last sequence update)	
DT	01-JUN-2002	(TREMBlrel. 21, Last annotation update)	
DE	Chymopasin (Chymotrypsin A CTRA-1).		
GN	CTRL OR CTRL1.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		

Q9ER05	PRELIMINARY;	PRT;	264 AA.
ID	Q9ER05		
AC	Q9ER05;		
DT	01-MAR-2001	(TREMBlrel. 16, Created)	
DT	01-MAR-2001	(TREMBlrel. 16, Last sequence update)	
DT	01-JUN-2002	(TREMBlrel. 21, Last annotation update)	
DE	Chymopasin (Chymotrypsin A CTRA-1).		
GN	CTRL OR CTRL1.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		

```

CC      TRYPsin FAMILY.
DR      EMBL; AK003060; BAB22539.1; -.
DR      HSSP; P00766; 1GCT.
DR      MEROPS; S01.152; -.
DR      MGD; MGI:1913723; 2200008D09Rik.
DR      InterPro; IPR001314; Chymotrypsin.
DR      InterPro; IPR001254; Ser_protease_Try.
DR      Pfam; PF00089; trypsin; 1.
DR      PRINTS; PR00722; CHYMOTRYPSIN.
DR      SMART; SM00020; Tryp_SPC; 1.
DR      PROSITE; PS50240; TRYPsin_DOM; 1.
DR      PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.
DR      PROSITE; PS00135; TRYPsin_SER; 1.
KW      Hydrolase; Serine protease.
SQ      SEQUENCE 263 AA; 27821 MW; 2620A27AFBA5D04D CRC64;

Query Match      42.8%; Score 71; DB 11; Length 263;
Best Local Similarity 48.3%; Pred. No. 0.11;
Matches 14; Conservative 5; Mismatches 10; Indels 0; Gaps

QY      1 MLLSLTSLVLLGSSWGCVPPIAIPALS 29
DB      1 MAFWLWSCFALVGNFGCGVPPIQPVLT 29

RESULT 6
Q9DX8X      PRELIMINARY; PRT; 263 AA.
ID      Q9DX8X
AC      Q9DX8X;
DT      01-JUN-2001 (TrEMBLrel. 17, Created)
DD      01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT      01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE      2200008D09Rik protein.
GN      2200008D09Rik.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX      MEDLINE=21085660; PubMed=11217851;
RA      Kawai J., Shingagawa A., Shibata Y., Yoshino M., Itoh M., Ishii Y.,
RA      Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA      Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA      Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA      Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA      Fleischmann W., Gaasterland T., Glass C., King B., Kochiwa H.,
RA      Kuehl P., Lewis S., Matsuo Y., Nikola I., Pesole G., Quackenbush J.,
RA      Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA      Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA      Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA      Brownstein M.J., Bult C., Fietcher C., Fujita M., Gariboldi M.,
RA      Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA      Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA      Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA      Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA      Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA      Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kotsuki S.,
RA      Hayashizaki Y.;
RT      Functional annotation of a full-length mouse cDNA collection.*;
RL      Nature 409:685-690(2001).
CC      -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC      TRYPsin FAMILY.
DR      EMBL; AK007566; BAB25112.1; -.
DR      HSSP; P00766; 1GCT.
DR      MEROPS; S01.152; -.
DR      MGD; MGI:1913723; 2200008D09Rik.
DR      InterPro; IPR001314; Chymotrypsin.
DR      InterPro; IPR001254; Ser_protease_Try.
DR      Pfam; PF00089; trypsin; 1.
DR      PRINTS; PR00722; CHYMOTRYPSIN.
DR      SMART; SM00020; Tryp_SPC; 1.

```

SQ SEQUENCE 263 AA; 27822 MW; 28C4487AF1A26B27 CRC64;

RA GaT

Query Match 35.8%; Score 59.5; DB 13; Length 261;  
Best Local Similarity 48.5%; Pred. No. 3.8;  
Matches 16; Conservative 5; Mismatches 9; Indels 3; Gaps 2

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;  
OC Pleuronectoidei; Paralichthyidae; Parulichthys.  
OX NCBI\_TaxID=8255;  
RN [1]  
RC SEQUENCE FROM N.A.  
RP TISSUE=PANCREAS;  
RC Suzuki T., Srivastava A.S., Kurokawa T.;  
RT "Japanese flounder mRNA for chymotrypsinogen 2.";  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
CC TRYPSIN FAMILY.  
DR EMBL: AB029754; BAA82366.1; -.  
DR HSSP: P00766; 1CHG.  
DR MEROPS: S01.152; -.  
DR InterPro: IPR001314; Chymotrypsin.  
DR InterPro: IPRO01254; Ser\_protease\_Try.  
DR Pfam: PF00089; trypsin; 1.  
DR PRINTS: PR00722; CHYMOTRYPSIN.  
DR SMART: SM00020; Tryp\_SPC; 1.  
DR PROSITE: PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE: PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease.  
SQ SEQUENCE 260 AA; 27793 MW; 9F583044E22F78C0 CRC64;

Query Match 33.7%; Score 56; DB 13; Length 260;  
Best Local Similarity 45.0%; Pred. No. 11;  
Matches 9; Conservative 5; Mismatches 6; Indels 0; G

OY 10 LVLLGGSGCGVPAITPALS 29  
| |:::||||||| |:::  
DB 7 LAFAGAYGCGSPAIPPVIT 26

RESULT 12

ID	QBSBN3	PRELIMINARY;	PRT;	140 AA.
AC	QBSBN3;			
DT	01-JUN-2002 (TrEMBLrel. 21, Created)			
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)			
DE	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)			
DE	Gp21 protein.			
OS	bacteriophage phi3626.			
OC	Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.			
OX	NCBI_TaxID=190478;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	Zimmer M., Scherer S., Loessner M.J.;			
RT	"Bacteriophage phi3626 complete genome.";			
RL	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AY082070; AAL96791.1; .			
SQ	SEQUENCE 140 AA; 15944 MW; 8D395E570A0A153C CRC64;			

Query Match 33.4%; Score 55.5; DB 9; Length 140;  
Best Local Similarity 38.7%; Pred. No. 7.2;  
Matches 12; Conservative 8; Mismatches 10; Indels 1; G

OY 3 LLESLTSLVLGGS-SWGCGVPAITPALSYNQ 32  
::: ||::: || |||| | |:  
DB 4 ITAILLSLVFLVGSLWGCGDKRETTNLNQ 34

RESULT 13

ID	O96666	PRELIMINARY;	PRT;	441 AA.
AC	O96666;			
DT	01-MAY-1999 (TrEMBLrel. 10, Created)			
DT	01-MAY-1999 (TrEMBLrel. 10, Last sequence update)			
DE	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)			
DE	IMAGINAL disc growth factor 3.			
GN	IDGF3 OR CG4559.			

OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kawamura K., Shibata T., Peel D., Park S.-B., Bryant P.J.;  
 RT "A new family of growth factors produced by the fat body and active on  
 RL Drosophila imaginal disc cells.";  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF102238; AAC99419.1; -;  
 DR FlyBase: FBgn0020414; Idgf3.  
 DR InterPro: IPR001223; Glyco\_hydro\_18.  
 DR Pfam: PF00704; Glyco\_hydro\_18; 1.  
 DR ProDom: PD000471; Glyco\_hydro\_18; 1.  
 SQ SEQUENCE 441 AA; 49219 MW; 988C297D09CA0B08 CRC64;  
 Query Match 33.4%; Score 55.5; DB 5; Length 441;  
 Best Local Similarity 48.3%; Pred. No. 22;  
 Matches 14; Conservative 5; Mismatches 9; Indels 1; Gaps 1;  
 QY 2 LLLSLTSLVLLGGSGCGVPATPALSY 30  
 Db 210 LLLSLTV-LPNVNSWYDAPSIAPSLDF 237  
 RESULT 14  
 Q9V3B9 PRELIMINARY; PRT; 441 AA.  
 AC Q9V3B9;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE IDGF3 protein (GH07453P).  
 GN IDGF3 OR C64559.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA Sutton G.G., Wortman J.R., Richards M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.M., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RX MEDLINE=99403001; PubMed=10471707;  
 RA Ashburner M., Misra S., Roote J., Lewis S.E., Blazej R., Davis T.,  
 RA Butenhoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,  
 RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,  
 RA Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim E., Lee B.,  
 RA Lewis S., Li P., Lomotan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,  
 RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,  
 RA Sethi H., Snir E., Svirskas R., Wan K.H., Weinburg T., Zhang R.,  
 RA Zieran L.L., Rubin G.M.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
 RA Nunco J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,  
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AE003651; AAF53537.1; -;  
 DR EMBL: AE003416; AAF45018.1; -;  
 DR EMBL: AY047561; AAK77293.1; -;  
 DR FlyBase: FBgn0020414; Idgf3.  
 DR InterPro: IPR001223; Glyco\_hydro\_18.  
 DR Pfam: PF00704; Glyco\_hydro\_18; 1.  
 DR ProDom: PD000471; Glyco\_hydro\_18; 1.  
 SQ SEQUENCE 441 AA; 49254 MW; F1938469FALL17755 CRC64;  
 Query Match 33.4%; Score 55.5; DB 5; Length 441;  
 Best Local Similarity 48.3%; Pred. No. 22;  
 Matches 14; Conservative 5; Mismatches 9; Indels 1; Gaps 1;  
 QY 2 LLLSLTSLVLLGGSGCGVPATPALSY 30  
 Db 210 LLLSLTV-LPNVNSWYDAPSIAPSLDF 237  
 RESULT 15  
 Q97Z55 PRELIMINARY; PRT; 835 AA.  
 AC Q97Z55;  
 DT 01-OCT-2001 (TREMBLrel. 18, Created)  
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)



```

DE Hypothetical protein SSO1079.
GN SSO1079.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Awayez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL; AE006726; AAK41341.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 835 AA; 92357 MW; BD28B5418BAF7250 CRC64;

Query Match 33.4%; Score 55.5; DB 17; Length 835;
Best Local Similarity 44.8%; Pred. No. 40;
Matches 13; Conservative 5; Mismatches 4; Indels 7; Gaps 1;

QY 1 MLLSLTSLV-----LLGSSWGCGVP 22
   :|||||:::|||||
Db 9 LLLSLTLPPIPVNSQSTVVISWGWTGTP 37

```

Search completed: December 20, 2002, 15:18:56  
Job time : 6.08696 secs



Result No.	Score	Query		ID	Description
		Match	Length		
1	150	90.4	264	CTRL_HUMAN	P40313 homo sapien
2	76	45.8	263	CTR2_CANFA	P04813 canis famil
3	73	44.0	263	CTR2_HUMAN	P17538 homo sapien
4	67	40.4	263	CTRB_RAT	P07338 rattus norv
5	58	34.9	312	OAC1_HUMAN	Q96kk4 homo sapien
6	55	33.1	232	FL3L_MOUSE	P49772 mus musculu
7	55	33.1	242	MTGA_KLEPN	Q48465 kiebistiell
8	55	33.1	268	CLCR_HUMAN	Q99895 homo sapien
9	52	31.3	219	ISPD_CHLFR	O84468 chlamydia t
10	51	30.7	484	DF5L_HUMAN	P57764 homo sapien
11	50	30.1	247	MCT5_MOUSE	P21844 mus musculu
12	50	30.1	363	CTB_TRYBB	P00164 trypanosoma
13	50	30.1	742	KM65_YEAST	Q3656 saccharomyc
14	49.5	29.8	312	US10_HSV1	P06486 herpes simp
15	49	29.5	119	VTU3_DROME	Q06521 drosophila
16	49	29.5	222	CP19_SOLTU	Q00652 solanum tub
17	49	29.5	344	YD33_MYCTU	Q10644 mycobacteri
18	48.5	29.2	65	MAXX_KLEPN	P29852 kiebistiella
19	48.5	29.2	310	PIP_SULSO	Q97ua2 sulfolobus
20	48.5	29.2	541	GT10_HUMAN	Q95528 homo sapien
21	48.5	29.2	606	MM17_HUMAN	Q9u129 homo sapien
22	48	28.9	245	CTRA_BOVIN	P00766 bos taurus
23	48	28.9	245	CTRB_BOVLN	P00767 bos taurus
24	48	28.9	263	CTRA_GADMO	P47796 gadus mornu
25	48	28.9	304	YG78_PSEAE	Q91347 pseudomonas
26	48	28.9	418	YK31_YEAST	P36142 saccharomyc
27	48	28.9	650	DMWD_MOUSE	Q08274 mus musculu
28	47.5	28.6	142	YOGA_BACSU	P54484 bacillus su
29	47.5	28.6	180	PAPF_HUMAN	P09466 homo sapien
30	47.5	28.6	218	ISPD_CHLMU	Q9pjti chlamydia m
31	47	28.3	116	NU3W_STRCA	O79102 struthio ca
32	47	28.3	413	MOEA_SALTY	Q56066 salmonella
33	47	28.3	568	DNL1_METAC	Q8tmt1 methanosarc

```

hum. MOL. GENET. 2:1389-1393(1993)).
-|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X71874; CAA50710.1; -
EMBL; X71877; CAA50711.1; -
HSP; P00763; LDPO.
MEROPS; S01.256; -
Genew; HGNC:2524; CTRL.
MIM; 118888; -
InterPro; IPR001314; Chymotrypsin.
InterPro; IPR001254; Ser.protease_Try.
Pfam; PF00089; trypsin; 1.
PRINTS; P000722; CHYMOTRYPSIN.
SMART; SM00020; Tryp_Spc; 1.
PROSITE; PS0240; TRYPSIN_DOM; 1.
PROSITE; PS00134; TRYPSIN_HIS; 1.
PROSITE; PS00135; TRYPSIN_SER; 1.
Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal.
SIGNAL 1 18 POTENTIAL.
FT PROPP 19 33 ACTIVATION PEPTIDE (POTENTIAL).
FT CHAIN 34 264 CHYMOTRYPSIN-LIKE PROTEASE CTRL-1.
FT ACT_SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 121 121 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 214 214 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 114 114 N-LINKED (GLCNAC. .) (POTENTIAL).
FT DISULFID 19 141 BY SIMILARITY.
FT DISULFID 60 76 BY SIMILARITY.
FT DISULFID 155 220 BY SIMILARITY.
FT DISULFID 187 201 BY SIMILARITY.

```

```

Query Match          45.8%; Score 76; DB 1; Length 263;
Best Local Similarity 55.2%; Pred. No. 0.0085;
Matches 16; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY  1 MLLSLTSLVLLGSSWCGVPAITPALS 29
      | | | | | | | | | | | | | |
Db   1 MAFLLWLLSCFALLGTAFGCGVPAIQVLS 29

RESULT 3
CTRB_HUMAN
ID* CTRB_HUMAN      STANDARD;      PRT;      263 AA.
AC   P17538;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Chymotrypsinogen B precursor (EC 3.4.21.1).
GN   CTRB1 OR CTRB.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
[1]
RN   SEQUENCE FROM N.A.
RP   TISSUE=Pancreas;
RC   MEDLINE=89134264; PubMed=2917002;
RA   Tomita N., Izumoto Y., Horii A., Doi S., Yokouchi H., Ogawa M.,
RA   Mori T., Matsubara K.;
RT   "Molecular cloning and nucleotide sequence of human pancreatic
RT   chymotrypsinogen cDNA.";
RL   Biochem. Biophys. Res. Commun. 158:569-575(1989).
[2]
RN   SEQUENCE FROM N.A.
RP   TISSUE=Pancreas;
RA   Strausberg R.;
RL   Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC   -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr|-Xaa, Trp|-Xaa,
CC   Phe|-Xaa, Leu|-Xaa.
CC   -1- SUBCELLULAR LOCATION: Extracellular.
CC   -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
-----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC   or send an email to license@isb-sib.ch).
-----
DR   EMBL; M24400; AAA52128.1; -.
DR   EMBL; BC005385; AAH05385.1; -.
DR   PIR; A31299; A31299.
DR   HSP; P00766; ICHG.
DR   MEROPS; S01.152; -.
DR   Genew; HGNC:2521; CTRB1.
DR   MTM; 118890; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001254; Ser_protease_Try.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPC; 1.
DR   PROSITE; PS00240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
FT   SIGNAL          1..18
FT   CHAIN           19..263      CHYMOTRYPSINOGEN B.
FT   CHAIN           19..31
FT   CHAIN           34..164      CHYMOTRYPSIN B, A CHAIN.
FT   CHAIN           167..263     CHYMOTRYPSIN B, B CHAIN.
FT   CHAIN           75..75       CHYMOTRYPSIN B, C CHAIN.
FT   ACT_SITE        75..75       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE        120..120     CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE        213..213     CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE        213..213     CHARGE RELAY SYSTEM (BY SIMILARITY).

```

FT DISULFID 19 140 BY SIMILARITY.  
FT DISULFID 60 76 BY SIMILARITY.  
FT DISULFID 154 219 BY SIMILARITY.  
FT DISULFID 186 200 BY SIMILARITY.  
FT DISULFID 209 238 BY SIMILARITY.  
SQ SEQUENCE 263 AA; 27870 MW; 4C1C055A49088701 CRC64;

Query Match 44.0%; Score 73; DB 1; Length 263;  
Best Local Similarity 55.2%; Pred. No. 0.021;  
Matches 16; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLLGSSWCGVPAITPALS 29  
| | | | | : : : : : | | | | |  
DB 1 MAFWLSCWALLGTTCGCVPAIHPVLS 29

RESULT 4  
CTRB\_RAT  
ID CTRB\_RAT STANDARD; PRT; 263 AA.  
AC P07338;  
DT 01-APR-1988 (Rel. 07, Created)  
DT 01-APR-1988 (Rel. 07, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Chymotrypsinogen B precursor (EC 3.4.21.1).  
GN CTRB1.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85054881; PubMed=6209274;  
RA Bell G I., Quinto C., Quiroga M., Valenzuela P., Craik C.S.,  
Rutter W.J.;  
RT Isolation and sequence of a rat chymotrypsin B gene.";  
RL J. Biol. Chem. 259:14265-14270(1984).  
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,  
Phe-I-Xaa, Leu-I-Xaa.  
CC -!- SUBCELLULAR LOCATION: Extracellular.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; K02298; AAA98732.1; -;  
PIR; A22658; KYRTP.  
DR HSP; P00766; ICHG.  
DR MEROPS; S01.152; -;  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR001254; Ser.protease\_Try.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; Tryp-SPC; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.  
FT SIGNAL 1 18  
FT CHAIN 19 263 CHYMOTRYPSINOGEN B.  
FT CHAIN 19 31 CHYMOTRYPSIN B, A CHAIN.  
FT CHAIN 34 164 CHYMOTRYPSIN B, B CHAIN.  
FT CHAIN 167 263 CHYMOTRYPSIN B, C CHAIN.  
FT ACT\_SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT DISULFID 19 140 BY SIMILARITY.  
FT DISULFID 60 76 BY SIMILARITY.  
FT DISULFID 154 219 BY SIMILARITY.

FT DISULFID 186 200 BY SIMILARITY.  
FT DISULFID 209 238 BY SIMILARITY.  
SQ SEQUENCE 263 AA; 27849 MW; ACAFD8AC8C8C64DA6D CRC64;  
Query Match 40.4%; Score 67; DB 1; Length 263;  
Best Local Similarity 44.8%; Pred. No. 0.12;  
Matches 13; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLLGSSWCGVPAITPALS 29  
| | | | | : : : : : | | | | |  
DB 1 MAFWLVSFCFALVGATFGCGVPTIOPVLT 29

RESULT 5  
OAC1\_HUMAN  
ID OAC1\_HUMAN STANDARD; PRT; 312 AA.  
AC Q96KK4;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Olfactory receptor 10C1 (Hs6M1-17).  
GN OR10C1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Whitaker H.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: PUTATIVE ODORANT RECEPTOR.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; AL035542; CAB44507.1; ALT\_INIT.  
DR InterPro; IPR000276; GPCR\_Rhodpsn.  
DR Pfam; PF00001; 7tm\_1; 1.  
DR PROSITE; PS00237; G-PROTEIN\_RECEP\_F1\_1; 1.  
DR PROSITE; PS00262; G-PROTEIN\_RECEP\_F1\_2; 1.  
KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
KW Multigene family; Olfaction.  
FT DOMAIN 1 24 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 25 48 1 (POTENTIAL).  
FT DOMAIN 49 56 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 57 78 2 (POTENTIAL).  
FT DOMAIN 79 99 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 100 119 3 (POTENTIAL).  
FT DOMAIN 120 138 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 139 157 4 (POTENTIAL).  
FT DOMAIN 158 194 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 195 218 5 (POTENTIAL).  
FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 236 258 6 (POTENTIAL).  
FT DOMAIN 259 271 7 (POTENTIAL).  
FT TRANSMEM 272 291 EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 292 312 CYTOPLASMIC (POTENTIAL).  
FT DISULFID 96 188 BY SIMILARITY.  
FT CARBOHYD 4 4 N-LINKED (GLCNAC... ) (POTENTIAL).  
SQ SEQUENCE 312 AA; 34351 MW; B6FD6E0E700CB2CC CRC64;

Query Match 34.9%; Score 58; DB 1; Length 312;  
Best Local Similarity 60.0%; Pred. No. 2.1;  
Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 LLLSLTSLVLLGSSWCGV 21

```
|||| : | | | | |
Db 133 LLLSHRVCLQAGSAMACGV 152

RESULT 6
FL3L_MOUSE
ID FL3L_MOUSE STANDARD; PRT; 232 AA.
AC P49772; Q64085;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE SL cytokine precursor (Fms-related tyrosine kinase 3 ligand) (Flt3
ligand) (Flt3L).
GN FLT3LG OR FLT3L.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
RP MEDLINE=94195428; PubMed=8145851;
RA Hannum C., Culpepper J., Campbell D., McClanahan T., Zurawski S.,
RA Bazan J.P., Kastelein R., Hudak S., Wagner J., Mattison J., Luh J.,
RA Duda G., Martina N., Peterson D., Menon S., Shanafelt A.,
RA Muench M., Kelnier G., Namikawa R., Rennick D., Roncarolo M.G.,
RA Zlotnik A., Rosnet O., Dubreuil P., Birnbaum D., Lee F.;
RT "Ligand for FLT3/FLK2 receptor tyrosine kinase regulates growth of
RT haematopoietic stem cells and is encoded by variant RNAs.";
RL Nature 368:643-648(1994).
[2]
SEQUENCE FROM N.A.
RP STRAIN=SJL/J;
RC MEDLINE=94084791; PubMed=7505204;
RA Lyman S.D., James L., Vandenbos T., Devries P., Brasel K.,
RA Gliniak B., Hollingsworth L.T., Picha K.S., McKenna H.J.,
RA Splet R.R., Fletcher F.A., Maraskovsky E., Farrah T.,
RA Foxworth D., Williams D.E., Beckmann M.P.;
RT "Molecular cloning of a ligand for the flt3/flk-2 tyrosine kinase
RT receptor: a proliferative factor for primitive hematopoietic cells.";
RL Cell 75:1157-1167(1993).
[3]
SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RP MEDLINE=96032581; PubMed=7566977;
RA Lyman S.D., Stocking K., Davidson B., Fletcher F., Johnson L.,
RA Escobar S.;
RT "Structural analysis of human and murine flt3 ligand genomic loci.";
RL Oncogene 11:1165-1172(1995).
[4]
SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RP MEDLINE=95124710; PubMed=7824267;
RA Lyman S.D., James L., Escobar S., Downey H., de Vries P.,
RA Brasel K., Stocking K., Beckmann M.P., Copeland N.G.,
RA Cleveland L.S.;
RT "Identification of soluble and membrane-bound isoforms of the murine
RT flt3 ligand generated by alternative splicing of mRNAs.";
RL Oncogene 10:149-157(1995).
[5]
SEQUENCE FROM N.A.
RP McClanahan T., Culpepper J., Campbell D., Wagner J.,
RA Franz-Bacon K., Mattison J., Tsai S., Luh J., Guimares M.J.,
RA Mattel M.-G., Rosnet O., Birnbaum D., Hannum C.;
RA Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: STIMULATES THE PROLIFERATION OF EARLY HEMATOPOIETIC
CC CELLS. SYNERGIZES WELL WITH A NUMBER OF OTHER COLONY STIMULATING
CC FACTORS AND INTERLEUKINS.
CC -1- SUBUNIT: Homodimer (soluble isoform) (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Two soluble
CC isoforms are also produced by alternative splicing. One of which,
CC isoform 3/E6, is biologically active, while the other, isoform
CC 4/E6delta16, is inactive.
CC -1- ALTERNATIVE PRODUCTS: 4 isoforms; 1/6C (shown here), 2/5H, 3/E6
CC and 4/E6delta16; are produced by alternative splicing.
-----

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; U04807; AAA18000.1; -
CC EMBL; L23636; AAA39436.1; -
CC EMBL; U29875; AAA90951.1; -
CC EMBL; U29875; AAA90952.1; -
CC EMBL; S76459; AAB33069.1; -
CC EMBL; S76461; AAB33070.1; -
CC EMBL; S76464; AAB33071.1; -
CC EMBL; U44024; AAA93307.1; -
CC EMBL; U44024; AAA93308.1; -
CC EMBL; MG1:95560; Flt3L.
CC MGD; MGI:95560; Flt3L.
CC DR InterPro: IPR004213; Flt3_lig; 1.
CC DR Pfam: PF02947; flt3_lig; 1.
KW Cytokine; Glycoprotein; Transmembrane; Alternative splicing; Signal.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 232 SL CYTOKINE.
FT DOMAIN 27 189 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 190 210 POTENTIAL.
FT DOMAIN 211 232 CYTOPLASMIC (POTENTIAL).
FT DISULFID 31 112 BY SIMILARITY.
FT DISULFID 71 156 BY SIMILARITY.
FT DISULFID 120 161 BY SIMILARITY.
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 164 232 DSSTLL -> GSHAG (IN ISOFORM 2).
FT VARSPLIC 164 169 DSSTLL -> GSHAG (IN ISOFORM 3).
FT VARSPLIC 170 232 VQQCP -> GLIMQARLTLSKQSSCLSVLSAGIT (IN
FT VARSPLIC 159 163 ISOFORM 4).
FT CONFLICT 141 141 A -> G (IN REF. 1).
FT CONFLICT 198 198 MISSING (IN REF. 2).
FT CONFLICT 232 232 A -> G (IN REF. 1).
SQ SEQUENCE 232 AA; 26141 MW; 3A3680D3CB69FBA6 CRC64;

Query Match 33.1%; Score 55; DB 1; Length 232;
Best Local Similarity 61.1%; Pred. No. 3.8;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLGSSWG 18
Db 193 LLLLLPLTLVLLAAWG 210

RESULT 7
MTGA_KLEPN STANDARD; PRT; 242 AA.
ID MTGA_KLEPN
AC Q48465;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Monofunctional biosynthetic peptidoglycan transglycosylase
DE (EC 2.4.2.-) (Monofunctional Tgase).
GN MTGA.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OX NCBI_TaxID=573;
[1]
SEQUENCE FROM N.A.
RP STRAIN=M5a1;
RX MEDLINE=96228711; PubMed=8830253;
RA Spratt B.G., Zhou J., Taylor M., Merrick M.J.;
RT "Monofunctional biosynthetic peptidoglycan transglycosylases.";
RL Mol. Microbiol. 19:639-640(1996).
```

CC -1- FUNCTION: CELL WALL FORMATION.  
CC -1- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: Membrane-associated (Potential).  
CC -1- SIMILARITY: TO THE N-TERMINAL OF BACTERIAL CLASS 1A PENICILLIN-BINDING PROTEINS.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: Z54198; CAA90903.1; -  
DR InterPro: IPR001264; GT\_51.  
DR Pfam: PF00912; Transglycosyl; 1.  
DR ProDom: PD001895; GT\_51; 1.  
KW Peptidoglycan synthesis; Cell wall; Transferase; Transmembrane.  
FT TRANSMEM 19 39 POTENTIAL.  
SQ SEQUENCE 242 AA; 27196 MW; A12CAB4A331ABA49 CRC64;  
  
Query Match 33.1%; Score 55; DB 1; Length 242;  
Best Local Similarity 58.8%; Pred. No. 4;  
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
  
Qy 17 WCGGVPAITPALSYNOR 33  
|| ||||| ||::||  
Db 92 WGFVPAIEKALAHNER 108  
  
RESULT 8  
CLCR\_HUMAN STANDARD; PRT; 268 AA.  
AC Q9895; Q9NUH5; O00765;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Caldecrin precursor (EC 3.4.21.2) (Chymotrypsin C).  
GN CTRC OR CLCR.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND VARIANT TRP-80.  
RC TISSUE=Pancreas;  
RX MEDLINE=96221265; PubMed=8635596;  
RA Tomomura A., Akiyama M., Itoh H., Yoshino I., Tomomura M., Nishii Y.,  
RA Nolkura T., Saheki T.;  
RT "Molecular cloning and expression of human caldecrin.";  
RL FEBS Lett. 386:26-28(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Coville G.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 17-268 FROM N.A.  
RC TISSUE=Pancreas;  
RA Szegoleit A.;  
RT "A human pancreatic chymotrypsin: biochemical and molecular characterization.";  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RN CHARACTERIZATION.  
RX MEDLINE=98207038; PubMed=9538241;  
RA Yoshino-Yasuda I., Kobayashi K., Akiyama M., Itoh H., Tomomura A.,  
RA Saheki T.;  
RT "Caldecrin is a novel-type serine protease expressed in pancreas, but its homologue, elastase IV, is an artifact during cloning derived from caldecrin gene.";  
RL J. Biochem. 123:546-554(1998).  
CC -1- FUNCTION: HAS CHYMOTRYPSIN-TYPE PROTEASE ACTIVITY AND HYPOCALCEMIC

CC ACTIVITY.  
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Leu-|-Xaa, Tyr-|-Xaa, Phe-|-Xaa, Met-|-Xaa, Trp-|-Xaa, Gln-|-Xaa, Asn-|-Xaa.  
CC -1- TISSUE SPECIFICITY: PANCREAS.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: S82198; AAB47104.2; ALT\_SEQ.  
DR EMBL: AL031283; CAB77355.1; -  
DR EMBL: Y13697; CAA74031.1; -  
DR HSP; P00766; ICHG.  
DR MEROPS: S01.157; -  
DR Genew: HGNC:2523; CTRC.  
DR MIM: 601405; -  
DR InterPro: IPR001314; Chymotrypsin.  
DR InterPro: IPR001254; Ser\_protease\_Try.  
DR Pfam: PF00089; trypsin; 1.  
DR PRINTS: PR00722; CHYMOTRYPSIN.  
DR SMART: SM00020; TRYD\_SPC; 1.  
DR PROSITE: PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE: PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal;  
KW Polymorphism.  
FT SIGNAL 1 16 POTENTIAL.  
FT PROPEP 17 29 ACTIVATION PEPTIDE.  
FT CHAIN 30 268 CALDECRIN.  
FT ACT\_SITE 74 74 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 121 121 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 216 216 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT DISULFID 17 141 BY SIMILARITY.  
FT DISULFID 59 75 BY SIMILARITY.  
FT DISULFID 155 222 BY SIMILARITY.  
FT DISULFID 186 202 BY SIMILARITY.  
FT DISULFID 212 243 BY SIMILARITY.  
FT CARBOHYD 25 25 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 52 52 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 226 226 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT VARIANT 80 80 R->W.  
FT CONFLICT 16 16 /FTID=VAR\_010928.  
FT CONFLICT 52 52 S->T (IN REF. 1).  
FT CONFLICT 52 52 N->D (IN REF. 3).  
SQ SEQUENCE 268 AA; 29484 MW; 460BF33BA496516F CRC64;  
  
Query Match 33.1%; Score 55; DB 1; Length 268;  
Best Local Similarity 40.7%; Pred. No. 4.3;  
Matches 11; Conservative 5; Mismatches 11; Indels 0; Gaps 0;  
  
Qy 3 LLSLTSLVLLSSWGCGVPAITPALS 29  
:|:|:|:|:|:|:|  
Db 1 MGIIVLAALLACASSCGVSPFPNLS 27  
  
RESULT 9  
ISPD\_CHLTR STANDARD; PRT; 219 AA.  
AC O84468;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE 2-C-methyl-D-erythritol 4-phosphate cytidyltransferase (EC 2.7.7.60)  
DE (4-diphosphocytidyl-2C-methyl-D-erythritol synthase) (MEP)  
DE cytidyltransferase) (MCT).  
GN ISPD OR CT462.  
OS Chlamydia trachomatis.  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.





Immunodetection with specific antipeptide Ig.;"

RT J. Immunol. 149:2466-2472(1992).

RL -!- SUBCELLULAR LOCATION: Secretory granules.

CC -!- TISSUE SPECIFICITY: MAST CELLS.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. GRANZYME SUBFAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

CC EMBL; X68805; CAA48705.1; ALT\_INIT.

DR EMBL; M73759; AAM40105.1; -.

DR EMBL; M73760; -; NOT\_ANNOTATED\_CDS.

DR EMBL; M68898; AAA39492.1; -.

DR PIR; A41076; A41076.

DR PIR; C35646; C35646.

DR HSSP; P23946; 1KLT.

DR MEROPS; S01150; -.

DR MGD; MGI:96941; Mcpt5.

DR InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR001254; Ser\_protease\_Try.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR SMART; SM00020; Tryp\_SPC; 1.

DR PROSITE; PS02040; TRYPsin\_DOM; 1.

DR PROSITE; PS00134; TRYPsin\_HIS; 1.

DR PROSITE; PS00135; TRYPsin\_SER; 1.

KW Hydrolase; Serine protease; Signal; Glycoprotein; Zymogen.

FT SIGNAL 1 19 POTENTIAL.

FT PROPEP 20 21 ACTIVATION PEPTIDE.

FT CHAIN 22 247 MAST CELL PROTEASE 5.

FT ACT\_SITE 66 66 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT\_SITE 110 110 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT\_SITE 203 203 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT DISULFID 51 67 BY SIMILARITY.

FT DISULFID 144 209 BY SIMILARITY.

FT DISULFID 175 188 BY SIMILARITY.

FT CARBOHYD 80 80 N-LINKED (GLCNAC... ) (POTENTIAL).

FT CONFLICT 5 5 T -> A (IN REF. 1).

FT CONFLICT 51 51 C -> R (IN REF. 4).

FT CONFLICT 224 224 A -> R (IN REF. 3).

SQ SEQUENCE 247 AA; 27586 MW; 24C290CF61237DC7 CRC64;

Query Match 30.1%; Score 50; DB 1; Length 247;

Best Local Similarity 39.5%; Pred. No. 18;

Matches 15; Conservative 5; Mismatches 10; Indels 8; Gaps 1;

OY 1 MLLSLTSLVLLGSSWCGG-----VPAITPALS 30

DB 1 MHLTLHLLLLLSSTKAGIEIGTECIPHSRPMAY 38

RESULT 12

CVB\_TRYBB

ID CVB\_TRYBB STANDARD; PRT; 363 AA.

AC P00164;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Cytochrome B.

GN COB OR CYTB.

OS Trypanosoma brucei brucei.

OG Mitochondrion.

OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.

OX NCBI\_TaxID=5702;

RN [1]

RP SEQUENCE FROM N.A. (KINETOPLAST).

RX MEDLINE=84041494; PubMed=6314266;

RA Benne R., de Vries B.F., van den Burg J., Klaver B.;

"The nucleotide sequence of a segment of Trypanosoma brucei

mitochondrial maxi-circle DNA that contains the gene for

apocytocrome b and some unusual unassigned reading frames.;"

RL Nucleic Acids Res. 11:6925-6941(1983).

CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE

CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A

CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL

CC COUPLED TO ATP SYNTHESIS.

CC -!- COPACITOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY

CC BOUND TO THE PROTEIN.

CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,

CC CYTOCHROME C1 AND THE RIESKE PROTEIN.

CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

CC EMBL; X00017; CAA24915.1; -.

DR PIR; A00160; CBUTB.

DR InterPro; IPR000179; Cyt\_b\_b6.

DR InterPro; IPR001230; Prenyl\_site.

DR Pfam; PF00032; cytochrome\_b\_c; 1.

DR Pfam; PF00033; cytochrome\_b\_n; 1.

DR PROSITE; PS00192; CYTOCHROME\_B\_HEME; 1.

DR PROSITE; PS00193; CYTOCHROME\_B\_OO; 1.

KW Electron transport; Mitochondrion; Respiratory chain; Transmembrane;

FT METAL 74 74 IRON 1 (HEME B562 AXIAL LIGAND).

FT METAL 88 88 IRON 2 (HEME B566 AXIAL LIGAND).

FT METAL 175 175 IRON 2 (HEME B562 AXIAL LIGAND).

FT METAL 189 189 IRON 1 (HEME B566 AXIAL LIGAND).

SQ SEQUENCE 363 AA; 43455 MW; 2A37CD827E55923B CRC64;

Query Match 30.1%; Score 50; DB 1; Length 363;

Best Local Similarity 46.4%; Pred. No. 25;

Matches 13; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

OY 3 LLSLTLVLLGSSWCGGVPATITPALS 30

DB 314 LLWLYSLILFYISWMSGFLALVYVLAY 341

RESULT 13

KM65\_YEAST

ID KM65\_YEAST STANDARD; PRT; 742 AA.

AC O03656;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Probable serine/threonine-protein kinase YMR216C (BC 2.7.1.-).

GN YMR216C OR YMR261.10C.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

OX NCBI\_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=S288C / AB972;

RA Dedman K., Brown D., Bowman S., Barrell B.G., Rajandream M.A.,

RA Walsh S.V.;

RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

CC SIMILAR TO S.POMBE DSK1.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: Z49809; CAA89931.1; -  
DR SGD: S0004829; YMR216C.  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR002290; Ser\_thr\_pkinase.  
DR Pfam: PF00069; pkinase; 1.  
DR ProDom: PD000001; Euk\_pkinase; 2.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE: PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR Hypothetical protein; Transferase; Serine/threonine-protein kinase;  
KW ATP-binding.  
FT DOMAIN 158 706 PROTEIN KINASE.  
FT NP\_BIND 164 172 ATP (BY SIMILARITY).  
FT BINDING 187 187 ATP (BY SIMILARITY).  
FT ACT\_SITE 294 294 BY SIMILARITY.  
SQ SEQUENCE 742 AA; 83237 MW; C775F10B30C950FC CRC64;  
  
Query Match 30.1%; Score 50; DB 1; Length 742;  
Best Local Similarity 80.0%; Pred. No. 49;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 11 VLLGSSWGCG 20  
IIII: IIII  
DB 575 VLLGAPWGCG 584  
  
RESULT 14  
US10\_HSV11  
ID US10\_HSV11 STANDARD; PRT; 312 AA.  
AC P06486;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 01-APR-1993 (Rel. 25, Last annotation update)  
DE Virion protein US10.  
GN US10.  
OS Herpes simplex virus (type 1 / strain 17).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Simplexvirus.  
OX NCBI\_TaxID=10299;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85160822; PubMed=2984429;  
RA McGeoch D.J., Dolan A., Donald S., Rixon F.J.;  
RT "Sequence determination and genetic content of the short unique  
RL region in the genome of herpes simplex virus type 1.";  
RL J. Mol. Biol. 181:1-13(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=84169548; PubMed=6324121;  
RA Rixon F.J., McGeoch D.J.;  
RT "A 3' co-terminal family of mRNAs from the herpes simplex virus type  
RT 1 short region: two overlapping reading frames encode unrelated  
RT polypeptide one of which has highly reiterated amino acid sequence.";  
RL Nucleic Acids Res. 12:2473-2487(1984).  
CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUP TOGETHER HSV-1 US10,  
CC EHV-1 66, EHV-4 ORF3, AND VZV 64/69.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: X14112; CAA32275.1; -  
DR EMBL: X02138; CAA26064.2; -  
DR EMBL: X00428; CAA45126.1; -  
DR EMBL: L00036; AAA96678.1; -

DR PIR: A05242; QOBE07.  
DR InterPro: IPR000087; Collagen.  
DR InterPro: IPR000714; EHV\_Unk.  
DR Pfam: PF02053; Gene66; 1.  
KW Zinc-finger. 271 293 POTENTIAL.  
FT 2N\_FING  
SQ SEQUENCE 312 AA; 34055 MW; 12D01B0E7C920EA3 CRC64;  
  
Query Match 29.8%; Score 49.5; DB 1; Length 312;  
Best Local Similarity 54.5%; Pred. No. 26;  
Matches 12; Conservative 3; Mismatches 6; Indels 1; Gaps 1;  
  
QY 7 TSLVLGSSWGCGVPAITPAL 28  
IIII: IIII  
DB 291 THCLHLFG-AFGGDPALTPPL 311  
  
RESULT 15  
VTU3\_DROME  
ID VTU3\_DROME STANDARD; PRT; 119 AA.  
AC Q06521;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Vitelline membrane protein Vm34Ca precursor.  
GN Vm34CA OR Vm34C.1.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94123988; PubMed=8293994;  
RA Scherer L.J., Harris D.H., White M.K., Steel L.S., Jin J.,  
RA Petri W.H.;  
RT "Comparative analysis of the sequence and structure of two Drosophila  
RL melanogaster genes encoding vitelline membrane proteins.";  
RL Gene 136:121-127(1993).  
RN [2]  
RP SEQUENCE OF 1-96 FROM N.A.  
RC STRAIN=Daekwanryeong;  
RX MEDLINE=85257433; PubMed=3926479;  
RA Mindrinos M.N., Scherer L.J., Garcini F.J., Kwan H., Jacobs K.A.,  
RA Petri W.H.;  
RT "Isolation and chromosomal location of putative vitelline membrane  
RT genes in Drosophila melanogaster.";  
RL EMBO J. 4:147-153(1985).  
RN [3]  
RP SEQUENCE OF 69-106 FROM N.A.  
RC STRAIN=Oregon-R;  
RX MEDLINE=89065296; PubMed=3143615;  
RA Scherer L.J., Harris D.H., Petri W.H.;  
RT "Drosophila vitelline membrane genes contain a 114 base pair region  
RT of highly conserved coding sequence.";  
RL Dev. Biol. 130:786-788(1988).  
CC -1- FUNCTION: MAJOR EARLY EGG-SHELL PROTEIN.  
CC -1- TISSUE SPECIFICITY: FOLLICLE CELLS.  
CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING VITELLINE MEMBRANE  
CC BIOSYNTHESIS.  
CC -1- SIMILARITY: A 38 AMINO ACIDS REGION (VM DOMAIN) IS CONSERVED IN  
CC DROSOPHILA VITELLINE MEMBRANE PROTEINS.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: L08852; AAC37200.1; -  
DR EMBL: X01802; CAA25933.1; -

DR	EMBL; M22700; AAA29018.1; -.
DR	FlyBase; FBgn0003983; Vm34Ca.
KW	Signal; Structural protein; Eggshell.
FT	SIGNAL 1 19
FT	CHAIN 20 119
FT	DOMAIN 69 106
FT	CONFLICT 96 96
FT	SEQUENCE 119 AA; 11934 MW; 2F0629BE52005BFC CRC64;
SO	VITELLINE MEMBRANE PROTEIN VM34CA.

Query Match 29.5%; Score 49; DB 1; Length 119;  
Best Local Similarity 54.3%; Pred. No. 12;  
Matches 12; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

```
QY 12 LLGSSW--GCGVPAITPALSYN 31
      :|||: | | | | | | | | | |
DB 25 MLGSSYGGGYGKPPAAAPAPSYS 46
```

Search completed: December 20, 2002, 15:17:14  
Job time : 2.13043 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: December 20, 2002, 15:16:49 ; Search time 8.34783 Seconds  
(without alignments)  
530.147 Million cell updates/sec

Title: US-09-856-319B-4

Perfect score: 1418

Sequence: 1 MLLSLTLISLVLLGSSWCGG.....AMVTRVSKFSTWVQVWVAYN 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 106657 seqs, 16763532 residues

Total number of hits satisfying chosen parameters: 106657

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database : Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pdb.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pdb.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pdb.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pdb.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pdb.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pdb.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pdb.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pdb.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pdb.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pdb.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pdb.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pdb.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pdb.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pdb.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	889	62.7	192	10	US-09-925-297-529
2	822.5	58.0	263	10	US-09-888-615-96
3	485.5	34.2	812	9	US-09-335-325-1
4	485.5	34.2	812	10	US-09-788-142-1
5	485.5	34.2	812	10	US-09-761-120-1
6	485.5	34.2	812	10	US-09-873-676-81
7	484.5	34.2	270	10	US-09-923-779-152
8	484.5	34.2	273	10	US-09-925-297-695
9	484	34.1	791	9	US-09-967-386-1
10	484	34.1	810	10	US-09-946-893-2
11	468.5	33.0	453	9	US-09-978-295A-69
12	468.5	33.0	453	9	US-09-978-697-69
13	468.5	33.0	453	9	US-09-978-192A-69
14	468.5	33.0	453	12	US-10-052-586-64
15	468	33.0	248	10	US-09-925-301-1017
16	468	33.0	327	10	US-09-804-156-16
17	468	33.0	327	10	US-09-946-633-8
18	468	33.0	454	10	US-09-888-615-103
19	464.5	32.8	251	9	US-09-961-721-4

20	462.5	32.6	343	10	US-09-948-094-2	Sequence 2, Appli
21	457.5	32.3	249	9	US-09-961-721-5	Sequence 5, Appli
22	443.5	31.3	226	9	US-09-961-721-6	Sequence 6, Appli
23	443.5	31.3	226	9	US-10-045-367A-5	Sequence 5, Appli
24	443.5	31.3	226	10	US-09-910-151-6	Sequence 6, Appli
25	440.5	31.1	1169	9	US-09-870-759-126	Sequence 126, App
26	440	31.0	393	9	US-10-012-896-934	Sequence 934, App
27	440	31.0	393	10	US-09-759-143-934	Sequence 934, App
28	440	31.0	393	10	US-09-780-669-934	Sequence 934, App
29	440	31.0	393	10	US-09-822-827-934	Sequence 934, App
30	440	31.0	492	9	US-10-012-896-932	Sequence 932, App
31	440	31.0	492	10	US-09-759-143-932	Sequence 932, App
32	440	31.0	492	10	US-09-780-669-932	Sequence 932, App
33	440	31.0	492	10	US-09-822-827-932	Sequence 932, App
34	438	30.9	492	9	US-10-012-896-895	Sequence 895, App
35	438	30.9	492	10	US-09-759-143-895	Sequence 895, App
36	438	30.9	492	10	US-09-780-669-895	Sequence 895, App
37	438	30.9	492	10	US-09-879-792-14	Sequence 14, Appl
38	438	30.9	492	10	US-09-822-827-895	Sequence 895, App
39	437	30.8	384	9	US-09-981-353-23	Sequence 23, Appl
40	435.5	30.7	970	10	US-09-888-615-101	Sequence 101, App
41	434.5	30.6	320	10	US-09-888-615-90	Sequence 90, Appl
42	431.5	30.4	284	12	US-10-041-400A-7	Sequence 7, Appli
43	431.5	30.4	284	12	US-10-041-264A-7	Sequence 7, Appli
44	431.5	30.4	284	12	US-10-042-091A-7	Sequence 7, Appli
45	431	30.4	283	10	US-09-988-975A-1	Sequence 1, Appli

## ALIGNMENTS

RESULT 1  
US-09-925-297-529  
; Sequence 529, Application US/09925297  
; Patent No. US20020081659A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA105  
; CURRENT APPLICATION NUMBER: US/09/925,297  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05989  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 928  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 529  
; LENGTH: 192  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-925-297-529

Query Match 62.7%; Score 889; DB 10; Length 192;  
Best Local Similarity 84.7%; Pred. No. 1e-71;  
Matches 160; Conservative 19; Mismatches 10; Indels 0; Gaps 0;  
QY 5 SUTLSLVLGSSGCGVPAITPALSYNQIRIVGENAVPGSWPQVSLQDNTGPHFCGSL 64  
|||||SUTLSLVLGSSGCGVPAITPALSYNQIRIVGENAVPGSWPQVSLQDNTGPHFCGSL 64  
Db 1 SUTLSLVLGSSGCGVPAITPALSYNQIRIVGENAVPGSWPQVSLQDNTGPHFCGSL 60  
QY 65 ISPNVVVTAHCOVTPGRHFVVLGEYDRSSNAEPQVLSIARAITHPNNANTMNDLTL 124  
|||SISQSWVTAHCOVTPGRHFVVLGEYDRSSNAEPQVLSIARAITHPNNANTMNDLTL 120  
Db 61 ISQSWVTAHCOVTPGRHFVVLGEYDRSSNAEPQVLSIARAITHPNNANTMNDLTL 120  
QY 125 LKLASPARVTAQVSPVCLASTNEALPSGLTCVTTGGRISGVGNVTPARLQVVLPLVTV 184  
|||||LKLASPARVTAQVSPVCLASTNEALPSGLTCVTTGGRISGVGNVTPARLQVVLPLVTV 184  
Db 121 LKLASPAQVTRISPVCLASSNEALTEGLTCVTTGGRISGVGNVTPARLQVVLPLVTV 180  
QY 185 NQCROYWGA 193  
|||||NQCROYWGA 193  
Db 181 NQCROYWGS 189



COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/788.142  
FILING DATE: 16-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/866,735  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren, William L.  
REGISTRATION NUMBER: 36,714  
REFERENCE/DOCKET NUMBER: 05940-0129  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404) 818-3700  
TELEFAX: (404) 818-3799  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 812 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Murine  
IMMEDIATE SOURCE:  
CLONE: Plasminogen  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-761-120-1

Query Match 34.2%; Score 485.5; DB 10; Length 812;  
Best Local Similarity 40.0%; Pred. No. 2.9e-35;  
Matches 106; Conservative 35; Mismatches 107; Indels 17; Gaps 7;

Qy 8 LSLVLLGSSWCGVPAITPALSYNQRIVNGENAVPGSWPQVSLQDN-TGFHFCGSLIS 66  
Db 557 IPLCASASSFCGKQVPEPK-KCGRVVGCVANPHSPWQISLRTFTGQHFCGGTLIA 615  
Qy 67 PNWVYTAHC---QVTPGRHFVVLGEYDRSSNAEPVQVLSIARATHPNWNANTMNDLT 123  
Db 616 PEWLTAAHCLKSSRPEFYKVLGAHEEYIRGLDVQEISVAKLILEPN-----NRDIA 669  
Qy 124 LLKLSPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQVVLPLVT 183  
Db 670 LLKLSRPATITDKVIPACLPSPNYMVDRTICYITGWTGCTGFGA--GRLKEAQLPVIE 727  
Qy 184 VNQCR--QYWGARTIDAMICAG--GSGASSCGDGGPLVCOKGNTWVLIGIVSWGTKNC 239  
Db 728 NKVCNRVEYLNRRVKSTELCAGLAGVDSCQDGGPLVCFEKDKYILQGVTSWGLGCA 787  
Qy 240 NIQAPAMTRYSKFSTWINQVMAYN 264  
Db 788 RPNKPGVYVRVSRFVDWIEREMRN 812

RESULT 5

US-09-761-120-1

Sequence 1, Application US/09761120  
Patent No. US2002037847A1  
GENERAL INFORMATION:  
APPLICANT: Folkman, M. Judah  
APPLICANT: O'Reilly, Michael  
TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasminogen  
FILE REFERENCE: 05940-0151 (43171-252068)  
CURRENT APPLICATION NUMBER: US/09/761.120  
CURRENT FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: 09/309,821

Query Match 34.2%; Score 485.5; DB 10; Length 812;  
Best Local Similarity 40.0%; Pred. No. 2.9e-35;  
Matches 106; Conservative 35; Mismatches 107; Indels 17; Gaps 7;

Qy 8 LSLVLLGSSWCGVPAITPALSYNQRIVNGENAVPGSWPQVSLQDN-TGFHFCGSLIS 66  
Db 557 IPLCASASSFCGKQVPEPK-KCGRVVGCVANPHSPWQISLRTFTGQHFCGGTLIA 615  
Qy 67 PNWVYTAHC---QVTPGRHFVVLGEYDRSSNAEPVQVLSIARATHPNWNANTMNDLT 123  
Db 616 PEWLTAAHCLKSSRPEFYKVLGAHEEYIRGLDVQEISVAKLILEPN-----NRDIA 669  
Qy 124 LLKLSPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQVVLPLVT 183  
Db 670 LLKLSRPATITDKVIPACLPSPNYMVDRTICYITGWTGCTGFGA--GRLKEAQLPVIE 727  
Qy 184 VNQCR--QYWGARTIDAMICAG--GSGASSCGDGGPLVCOKGNTWVLIGIVSWGTKNC 239  
Db 728 NKVCNRVEYLNRRVKSTELCAGLAGVDSCQDGGPLVCFEKDKYILQGVTSWGLGCA 787  
Qy 240 NIQAPAMTRYSKFSTWINQVMAYN 264  
Db 788 RPNKPGVYVRVSRFVDWIEREMRN 812

RESULT 5

US-09-761-120-1

Sequence 1, Application US/09761120  
Patent No. US2002037847A1  
GENERAL INFORMATION:  
APPLICANT: Folkman, M. Judah  
APPLICANT: O'Reilly, Michael  
TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasminogen  
FILE REFERENCE: 05940-0151 (43171-252068)  
CURRENT APPLICATION NUMBER: US/09/761.120  
CURRENT FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: 09/309,821

PRIOR FILING DATE: 1999-05-11  
PRIOR APPLICATION NUMBER: 08/866,735  
PRIOR FILING DATE: 1997-05-30  
NUMBER OF SEQ ID NOS: 47  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 1  
LENGTH: 812  
TYPE: PRT  
ORGANISM: Murinae gen. sp.  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: Plasminogen  
US-09-761-120-1

Query Match 34.2%; Score 485.5; DB 10; Length 812;  
Best Local Similarity 40.0%; Pred. No. 2.9e-35;  
Matches 106; Conservative 35; Mismatches 107; Indels 17; Gaps 7;

Qy 8 LSLVLLGSSWCGVPAITPALSYNQRIVNGENAVPGSWPQVSLQDN-TGFHFCGSLIS 66  
Db 557 IPLCASASSFCGKQVPEPK-KCGRVVGCVANPHSPWQISLRTFTGQHFCGGTLIA 615  
Qy 67 PNWVYTAHC---QVTPGRHFVVLGEYDRSSNAEPVQVLSIARATHPNWNANTMNDLT 123  
Db 616 PEWLTAAHCLKSSRPEFYKVLGAHEEYIRGLDVQEISVAKLILEPN-----NRDIA 669  
Qy 124 LLKLSPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQVVLPLVT 183  
Db 670 LLKLSRPATITDKVIPACLPSPNYMVDRTICYITGWTGCTGFGA--GRLKEAQLPVIE 727  
Qy 184 VNQCR--QYWGARTIDAMICAG--GSGASSCGDGGPLVCOKGNTWVLIGIVSWGTKNC 239  
Db 728 NKVCNRVEYLNRRVKSTELCAGLAGVDSCQDGGPLVCFEKDKYILQGVTSWGLGCA 787  
Qy 240 NIQAPAMTRYSKFSTWINQVMAYN 264  
Db 788 RPNKPGVYVRVSRFVDWIEREMRN 812

RESULT 6

US-09-873-676-81

Sequence 81, Application US/09873676  
Patent No. US2002007289A1  
GENERAL INFORMATION:  
APPLICANT: Macdonald, Nicholas J.  
APPLICANT: Sim, Kim L.  
TITLE OF INVENTION: Angiotensin and Endostatin Binding Proteins and Methods of Use  
FILE REFERENCE: 05213-0378 (43170-259333)  
CURRENT APPLICATION NUMBER: US/09/873.676  
CURRENT FILING DATE: 2001-06-04  
PRIOR APPLICATION NUMBER: US 60/209,065  
PRIOR FILING DATE: 2000-06-02  
PRIOR APPLICATION NUMBER: US 60/289,387  
PRIOR FILING DATE: 2001-05-08  
NUMBER OF SEQ ID NOS: 123  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 81  
LENGTH: 812  
TYPE: PRT  
ORGANISM: Murinae gen. sp.  
US-09-873-676-81

Query Match 34.2%; Score 485.5; DB 10; Length 812;  
Best Local Similarity 40.0%; Pred. No. 2.9e-35;  
Matches 106; Conservative 35; Mismatches 107; Indels 17; Gaps 7;

Qy 8 LSLVLLGSSWCGVPAITPALSYNQRIVNGENAVPGSWPQVSLQDN-TGFHFCGSLIS 66  
Db 557 IPLCASASSFCGKQVPEPK-KCGRVVGCVANPHSPWQISLRTFTGQHFCGGTLIA 615  
Qy 67 PNWVYTAHC---QVTPGRHFVVLGEYDRSSNAEPVQVLSIARATHPNWNANTMNDLT 123  
Db 616 PEWLTAAHCLKSSRPEFYKVLGAHEEYIRGLDVQEISVAKLILEPN-----NRDIA 669





Db	545	SFDCGKGFQVDPK--KCPORVVGCGVAPHSPHWPQVSLURTRGMHFCGGTILISPEWLTAAH	603
Qy	76	C---QVTPGRHVVYGLDYDRSSNAEP--VQVLSIARAITHPNWNANTMNNDLTLLKLASPA	131
Db	604	CLEKSPRPSSYKVIILGAH--QEVNLEPHVQIEVSRULEPT-----RKDIALLLKLSPPA	656
Qy	132	RYTAQNSPVLCASTNEALPSGLTCVTGTGWRISGVGNVTTPARLQVVVLPTVTVNQCQRY-	190
Db	657	VITDKVPIACLPSPNYVYVADTECFITGWGETGTGFGA--GLLKEAQLPVIENKVCNRYE	714
Qy	191	-WGARIITDMACAG--CGSASSCGQSGDGGPLVCQKNTWVLIGIVSWGTKNCNIOAPMY	247
Db	715	FLNCRVOSTELCAGHLAGGTDCQSGDGGPLVCFEKDKYILQGVTSWGLGCGARPKNPGVI	774
Qy	248	TRVSKFSTWINQVMAYN	264
Db	775	VRVSREVTWIEGVMRNN	791

RESULT 10  
 US-09-946-893-2  
 ; Sequence 2, Application US/09946893  
 ; Patent No. US20020072494A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yihai  
 ; TITLE OF INVENTION: Materials and methods relating to endothelial cell growth  
 ; FILE REFERENCE: Newbourn  
 ; CURRENT APPLICATION NUMBER: US/09/946,893  
 ; CURRENT FILING DATE: 2001-09-05  
 ; PRIOR APPLICATION NUMBER: US 60/230,893  
 ; PRIOR FILING DATE: 2000-09-05  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 810  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-946-893-2

Query Match	34.1%	Score 484;	DB 10;	Length 810;
Best Local Similarity	41.2%	Pred. No. 3.9e-35;		
Matches 106; Conservative	31;	Mismatches 102;	Indels 18;	Gaps 8;

Qy	16	SWGCGVPAITPALSYNQRIIVNGENAVPGSPWQVSVLODNTGPHFCGGSILSPNWWVTAAH	75
Db	564	SFDCGKQVEPK-KCPCRVVGGCVAPHSPWQVSVLRTFRGMHFCGGSILSPWVLTAAH	622
Qy	76	C---QVTPGHRHVVVLGVDYDRSSNAEP--VQVLSSTARAITHPNNWANTMNDLTLLKLASPA	131
Db	623	CLEKSPRPSSKYVILGAH-QEVNLEPHVQVIEVSRLEPT-----RDIALLLKLSSPA	675
Qy	132	RYTAQVSPVCLASTNEALPSGLTFCVTTGWGRISGVGNVTPARLQVVVLPLVTWNQCQRY-	190
Db	676	VITDKVIPACLPSPNVVYADRTCEFTTGWGETGTGCA--GLLEKAEQLPVIENKVCNRYE	733
Qy	191	-WGARIITDAMICAG--GSGASSCGDGGGGLVCQKQNTWVLIGVSWGTKNCNIQAPMY	247
Db	734	FLNCRVOSTELCAGHLAGGTDCQDGGGGLVCFEKDKYILQGVTSWGLGCARPKNPGVY	793
Qy	248	TRVSKFTSWINQVMAYN	264
Db	794	VRVSREVTWIEGVMRNN	810

RESULT 11  
US-09-378-295A-69  
: Sequence 69, Application US/09978295A  
: Patent No. US2002015006A1  
: GENERAL INFORMATION:  
: APPLICANT: ASHKENAZI, Avl  
: APPLICANT: Baker Kevin P.  
: APPLICANT: Botstein, David

: APPLICANT: Desnoyers, Luc  
 : APPLICANT: Eaton, Dan  
 : APPLICANT: Ferrara, Napoleon  
 : APPLICANT: Filvaroff, Ellen  
 : APPLICANT: Fong, Sherman  
 : APPLICANT: Gao, Wei-Qiang  
 : APPLICANT: Gerber, Hanspeter  
 : APPLICANT: Gerritsen, Mary E.  
 : APPLICANT: Goddard, Audrey  
 : APPLICANT: Godowski, Paul J.  
 : APPLICANT: Grimaldi, J. Christopher  
 : APPLICANT: Gurney, Austin L.  
 : APPLICANT: Hillan, Kenneth J.  
 : APPLICANT: Kijavlin, Ivar J.  
 : APPLICANT: Kuo, Sophia S.  
 : APPLICANT: Napier, Mary A.  
 : APPLICANT: Pan, James;  
 : APPLICANT: Paoni, Nicholas F.  
 : APPLICANT: Roy, Margaret Ann  
 : APPLICANT: Shelton, David L.  
 : APPLICANT: Stewart, Timothy A.  
 : APPLICANT: Tunas, Daniel  
 : APPLICANT: Williams, P. Mickey  
 : APPLICANT: Wood, William I.  
 : TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 : FILE REFERENCE: P2630PIC11  
 : CURRENT APPLICATION NUMBER: US/09/978,295A  
 : CURRENT FILING DATE: 2001-10-15  
 : PRIOR APPLICATION NUMBER: 09/918585  
 : PRIOR FILING DATE: 2001-07-30  
 : PRIOR APPLICATION NUMBER: 60/062250  
 : PRIOR FILING DATE: 1997-10-17  
 : PRIOR APPLICATION NUMBER: 60/064249  
 : PRIOR FILING DATE: 1997-11-03  
 : PRIOR APPLICATION NUMBER: 60/065311  
 : PRIOR FILING DATE: 1997-11-13  
 : PRIOR APPLICATION NUMBER: 60/066364  
 : PRIOR FILING DATE: 1997-11-21  
 : PRIOR APPLICATION NUMBER: 60/077450  
 : PRIOR FILING DATE: 1998-03-10  
 : PRIOR APPLICATION NUMBER: 60/077632  
 : PRIOR FILING DATE: 1998-03-11  
 : PRIOR APPLICATION NUMBER: 60/077641  
 : PRIOR FILING DATE: 1998-03-11  
 : PRIOR APPLICATION NUMBER: 60/077649  
 : PRIOR FILING DATE: 1998-03-11  
 : PRIOR APPLICATION NUMBER: 60/077791  
 : PRIOR FILING DATE: 1998-03-12  
 : PRIOR APPLICATION NUMBER: 60/078004  
 : PRIOR FILING DATE: 1998-03-13  
 : PRIOR APPLICATION NUMBER: 60/078886  
 : PRIOR FILING DATE: 1998-03-20  
 : PRIOR APPLICATION NUMBER: 60/078936  
 : PRIOR FILING DATE: 1998-03-20  
 : PRIOR APPLICATION NUMBER: 60/078910  
 : PRIOR FILING DATE: 1998-03-20  
 : PRIOR APPLICATION NUMBER: 60/078939  
 : PRIOR FILING DATE: 1998-03-20  
 : PRIOR APPLICATION NUMBER: 60/079294  
 : PRIOR FILING DATE: 1998-03-25  
 : PRIOR APPLICATION NUMBER: 60/079656  
 : PRIOR FILING DATE: 1998-03-26  
 : PRIOR APPLICATION NUMBER: 60/079664  
 : PRIOR FILING DATE: 1998-03-27  
 : PRIOR APPLICATION NUMBER: 60/079689  
 : PRIOR FILING DATE: 1998-03-27  
 : PRIOR APPLICATION NUMBER: 60/079663  
 : PRIOR FILING DATE: 1998-03-27  
 : PRIOR APPLICATION NUMBER: 60/079728  
 : PRIOR FILING DATE: 1998-03-27  
 : PRIOR APPLICATION NUMBER: 60/079786  
 : PRIOR FILING DATE: 1998-03-27

; PRIOR APPLICATION NUMBER: 60/079920  
 ; PRIOR FILING DATE: 1998-03-30  
 ; PRIOR APPLICATION NUMBER: 60/079923  
 ; PRIOR FILING DATE: 1998-03-30  
 ; PRIOR APPLICATION NUMBER: 60/080105  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080107  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080165  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080194  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080327  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/080328  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/080333  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/080334  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/081070  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081049  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081071  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081195  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081203  
 ; PRIOR FILING DATE: 1998-04-09  
 ; PRIOR APPLICATION NUMBER: 60/081229  
 ; PRIOR FILING DATE: 1998-04-09  
 ; PRIOR APPLICATION NUMBER: 60/081955  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/081817  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/081819  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/081952  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/081838  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/082568  
 ; PRIOR FILING DATE: 1998-04-21  
 ; PRIOR APPLICATION NUMBER: 60/082569  
 ; PRIOR FILING DATE: 1998-04-21  
 ; PRIOR APPLICATION NUMBER: 60/082704  
 ; PRIOR FILING DATE: 1998-04-22  
 ; PRIOR APPLICATION NUMBER: 60/082804  
 ; PRIOR FILING DATE: 1998-04-22  
 ; PRIOR APPLICATION NUMBER: 60/082700  
 ; PRIOR FILING DATE: 1998-04-22  
 ; PRIOR APPLICATION NUMBER: 60/082797  
 ; PRIOR FILING DATE: 1998-04-22  
 ; PRIOR APPLICATION NUMBER: 60/082796  
 ; PRIOR FILING DATE: 1998-04-23  
 ; PRIOR APPLICATION NUMBER: 60/083336  
 ; PRIOR FILING DATE: 1998-04-27  
 ; PRIOR APPLICATION NUMBER: 60/083322  
 ; PRIOR FILING DATE: 1998-04-28  
 ; PRIOR APPLICATION NUMBER: 60/083392  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083495  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083496  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083499  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083545  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083554  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083558

; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083559  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083500  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083742  
 ; PRIOR FILING DATE: 1998-04-30  
 ; PRIOR APPLICATION NUMBER: 60/084366  
 ; PRIOR FILING DATE: 1998-05-05  
 ; PRIOR APPLICATION NUMBER: 60/084414  
 ; PRIOR FILING DATE: 1998-05-06  
 ; PRIOR APPLICATION NUMBER: 60/084441  
 ; PRIOR FILING DATE: 1998-05-06  
 ; PRIOR APPLICATION NUMBER: 60/084637  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084639  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084640  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084598  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084600  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084627  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084643  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/085339  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085338  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085323  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085582  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085700  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085689  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085579  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085580  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085573  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085704  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085697

Query Match 33.0%; Score 468.5; DB 9; Length 453;

Best Local Similarity 38.1%; Pred. No. 4.5e-34;

Matches 91; Conservative 40; Mismatches 97; Indels 11; Gaps 6;

QY 30 YNORIVGENAVPGSWPQVSLQDNTGFHFCGSGSLISPNWVTTAAHC---QVTPGRHFVV 86  
 Db 213 YSSRIVGGNNMSLLSQWPMQASLQFQ-GYHLCGGSVITPLWIIITAAHCYDYLPLPKSWTIQ 271  
 QY 87 LGEYDRSSNAEPYQVLSIARAITHPNWNANTMNDLTLKLSAPRYTAOVSPVCLASTN 146  
 Db 272 VGLVSLLDNPPASHL--VEKIVYHSKYKPRRLGNLALMKLAGPLTFENEMIQVCLPNS 329  
 QY 147 EALPSGLTCVTTGWRISGVNVTTPARLQOVVLPLVTVNOC--ROYWGARITDAMICAG- 203  
 Db 330 ENFPDGKVCWTSCWGATEDGGDASPV-LNHAAPVLSIKNCINHRDVGYYGIIISPSMLCAGY 388  
 QY 204 -GSCASSCCGDSGGPLVCQKNTWVLIGIVSWGTKNCNICQAPAMYTFVSKFSTWQVNM 261  
 Db 389 LTGGVDSOCDGSGGPLVCQERRLKLVGATSGFGICAEVKNKPGYVTRVTSFLDWIHEQM 447

RESULT 12

US-09-978-697-69

; Sequence 69, Application US/09978697

Patent No. US20020169284A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Baker Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan  
APPLICANT: Ferrara, Napoleon  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kijavlin, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2630P1C27  
CURRENT APPLICATION NUMBER: US/09/978,697  
CURRENT FILING DATE: 2001-10-16  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29

; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 33.0%; Score 468.5; DB 9; Length 453;
Best Local Similarity 38.1%; Pred. No. 4.5e-34;
Matches 91; Conservative 40; Mismatches 97; Indels 11; Gaps 6;

QY 30 YNORIVGENAVPGSPQVSLQDNTGFHFCGSLISPNWVYTAHC--QVTPGRHFVY 86
DB 213 YSSRIVGNNLSLSPQWQASQFQ-GYHLCCGGSVITPLWITAAHCYVDLYLPKSWTIQ 271
QY 87 LGRYDRSSNAEPQVLSIARATHPNWNNANTMNDLTLKLSAPRYTAQVSPVCLASTN 146
DB 272 VGLVSLLDNAPSHL--VEKIVYHSYKPKRLGNDALMKLAGPLTFNEMIQPVCLPNSF 329
QY 147 EALPSGLTCVTTGWGRISGVNGVTPARLQOVVLPVTVNQC--ROYWGARTDAMTCAG- 203
DB 330 ENFPDGKVCWTSGWATEDGGDASPY-LNHAAPVPLISNKNRHDVYGGIISPSMLCAGY 388
QY 204 -GSGASSCCDGGPLVCGKNGTWNVLIGVSGTKNCNTQAPAMYTRVSKFTWVNOVM 261
DB 389 LTGVDVSCQDGGPLVCGQERRLWLKLVGATSFSGCAEVNKPGVYTRVTSFLDWIHEQM 447

RESULT 13
US-09-978-192A-69
; Sequence 69, Application US/09978192A
; Patent No. US2002017753A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630PIC9
; CURRENT APPLICATION NUMBER: US/09/978, 192A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26

✓	PRIOR FILING DATE:	1998-04-29
✓	PRIOR APPLICATION NUMBER:	60/083496
✓	PRIOR FILING DATE:	1998-04-29
✓	PRIOR APPLICATION NUMBER:	60/083499
✓	PRIOR FILING DATE:	1998-04-29
✓	PRIOR APPLICATION NUMBER:	60/083545
✓	PRIOR FILING DATE:	1998-04-29
✓	PRIOR APPLICATION NUMBER:	60/083554
✓	PRIOR FILING DATE:	1998-04-29
✓	PRIOR APPLICATION NUMBER:	60/083558
✓	PRIOR FILING DATE:	1998-04-29
✓	PRIOR APPLICATION NUMBER:	60/083559
✓	PRIOR FILING DATE:	1998-04-29
✓	PRIOR APPLICATION NUMBER:	60/083500
✓	PRIOR FILING DATE:	1998-04-29
✓	PRIOR APPLICATION NUMBER:	60/083742
✓	PRIOR FILING DATE:	1998-04-30
✓	PRIOR APPLICATION NUMBER:	60/084366
✓	PRIOR FILING DATE:	1998-05-05
✓	PRIOR APPLICATION NUMBER:	60/084414
✓	PRIOR FILING DATE:	1998-05-06
✓	PRIOR APPLICATION NUMBER:	60/084441
✓	PRIOR FILING DATE:	1998-05-06
✓	PRIOR APPLICATION NUMBER:	60/084637
✓	PRIOR FILING DATE:	1998-05-07
✓	PRIOR APPLICATION NUMBER:	60/084639
✓	PRIOR FILING DATE:	1998-05-07
✓	PRIOR APPLICATION NUMBER:	60/084640
✓	PRIOR FILING DATE:	1998-05-07
✓	PRIOR APPLICATION NUMBER:	60/084598
✓	PRIOR FILING DATE:	1998-05-07
✓	PRIOR APPLICATION NUMBER:	60/084600
✓	PRIOR FILING DATE:	1998-05-07
✓	PRIOR APPLICATION NUMBER:	60/084627
✓	PRIOR FILING DATE:	1998-05-07
✓	PRIOR APPLICATION NUMBER:	60/084643
✓	PRIOR FILING DATE:	1998-05-07
✓	PRIOR APPLICATION NUMBER:	60/085339
✓	PRIOR FILING DATE:	1998-05-13
✓	PRIOR APPLICATION NUMBER:	60/085338
✓	PRIOR FILING DATE:	1998-05-13
✓	PRIOR APPLICATION NUMBER:	60/085323
✓	PRIOR FILING DATE:	1998-05-13
✓	PRIOR APPLICATION NUMBER:	60/085582
✓	PRIOR FILING DATE:	1998-05-15
✓	PRIOR APPLICATION NUMBER:	60/085700
✓	PRIOR FILING DATE:	1998-05-15
✓	PRIOR APPLICATION NUMBER:	60/085689
✓	PRIOR FILING DATE:	1998-05-15
✓	PRIOR APPLICATION NUMBER:	60/085797
✓	PRIOR FILING DATE:	1998-05-15
✓	PRIOR APPLICATION NUMBER:	60/085580
✓	PRIOR FILING DATE:	1998-05-15
✓	PRIOR APPLICATION NUMBER:	60/085573
✓	PRIOR FILING DATE:	1998-05-15
✓	PRIOR APPLICATION NUMBER:	60/085704
✓	PRIOR FILING DATE:	1998-05-15
✓	PRIOR APPLICATION NUMBER:	60/085697

9

[illegible]

Db	330	ENPPDKVCTSWGATEDGGDASPV-LNHAAVPLISNWKICHNRDVGGLISPSMLCAGY	388
Qy	204	-GSGASSCQDGGPLVQCQGNWTWVLIGVSGWTGNKCNIQAPAMTRYRKSFTWNOVM	261
Db	389	LTGCVDSQCQDGGPLVQCQERRLWKLVGATSFEGICGAENVKPGVTVRTVTSFLDWIHEQM	447
RESULT 14			
US-10-052-586-64	US-10-052-586-64		
Sequence 64, Application US/10052586	Sequence 64, Application US/10052586		
Patent No. US20020127584A1	Patent No. US20020127584A1		
GENERAL INFORMATION:	GENERAL INFORMATION:		
APPLICANT: Baker, Kevin P.	APPLICANT: Baker, Kevin P.		
APPLICANT: Chen, Jian	APPLICANT: Chen, Jian		
APPLICANT: Desnoyers, Luc	APPLICANT: Desnoyers, Luc		
APPLICANT: Goddard, Audrey	APPLICANT: Goddard, Audrey		
APPLICANT: Godowski, Paul J.	APPLICANT: Godowski, Paul J.		
APPLICANT: Gurney, Austin L.	APPLICANT: Gurney, Austin L.		
APPLICANT: Pan, James	APPLICANT: Pan, James		
APPLICANT: Smith, Victoria	APPLICANT: Smith, Victoria		
APPLICANT: Watanabe, Colin K.	APPLICANT: Watanabe, Colin K.		
APPLICANT: Wood, William I.	APPLICANT: Wood, William I.		
APPLICANT: Zhang, Zemin	APPLICANT: Zhang, Zemin		
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC	TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC		
TITLE OF INVENTION: ACIDS ENCODING THE SAME	TITLE OF INVENTION: ACIDS ENCODING THE SAME		
FILE REFERENCE: P3430RIC1	FILE REFERENCE: P3430RIC1		
CURRENT APPLICATION NUMBER: US/10/052,586	CURRENT APPLICATION NUMBER: US/10/052,586		
CURRENT FILING DATE: 2002-01-15	CURRENT FILING DATE: 2002-01-15		
PRIOR APPLICATION NUMBER: 60/059263	PRIOR APPLICATION NUMBER: 60/059263		
PRIOR FILING DATE: 1997-09-18	PRIOR FILING DATE: 1997-09-18		
PRIOR APPLICATION NUMBER: 60/059266	PRIOR APPLICATION NUMBER: 60/059266		
PRIOR FILING DATE: 1997-09-18	PRIOR FILING DATE: 1997-09-18		
PRIOR APPLICATION NUMBER: 60/062250	PRIOR APPLICATION NUMBER: 60/062250		
PRIOR FILING DATE: 1997-10-17	PRIOR FILING DATE: 1997-10-17		
PRIOR APPLICATION NUMBER: 60/063120	PRIOR APPLICATION NUMBER: 60/063120		
PRIOR FILING DATE: 1997-10-24	PRIOR FILING DATE: 1997-10-24		
PRIOR APPLICATION NUMBER: 60/063121	PRIOR APPLICATION NUMBER: 60/063121		
PRIOR FILING DATE: 1997-10-24	PRIOR FILING DATE: 1997-10-24		
PRIOR APPLICATION NUMBER: 60/063486	PRIOR APPLICATION NUMBER: 60/063486		
PRIOR FILING DATE: 1997-10-21	PRIOR FILING DATE: 1997-10-21		
PRIOR APPLICATION NUMBER: 60/063540	PRIOR APPLICATION NUMBER: 60/063540		
PRIOR FILING DATE: 1997-10-28	PRIOR FILING DATE: 1997-10-28		
PRIOR APPLICATION NUMBER: 60/063541	PRIOR APPLICATION NUMBER: 60/063541		
PRIOR FILING DATE: 1997-10-28	PRIOR FILING DATE: 1997-10-28		
PRIOR APPLICATION NUMBER: 60/063544	PRIOR APPLICATION NUMBER: 60/063544		
PRIOR FILING DATE: 1997-10-28	PRIOR FILING DATE: 1997-10-28		
PRIOR APPLICATION NUMBER: 60/063564	PRIOR APPLICATION NUMBER: 60/063564		
PRIOR FILING DATE: 1997-10-28	PRIOR FILING DATE: 1997-10-28		
PRIOR APPLICATION NUMBER: 60/063734	PRIOR APPLICATION NUMBER: 60/063734		
PRIOR FILING DATE: 1997-10-29	PRIOR FILING DATE: 1997-10-29		
PRIOR APPLICATION NUMBER: 60/063870	PRIOR APPLICATION NUMBER: 60/063870		
PRIOR FILING DATE: 1997-10-31	PRIOR FILING DATE: 1997-10-31		
PRIOR APPLICATION NUMBER: 60/064103	PRIOR APPLICATION NUMBER: 60/064103		
PRIOR FILING DATE: 1997-10-31	PRIOR FILING DATE: 1997-10-31		
PRIOR APPLICATION NUMBER: 60/065311	PRIOR APPLICATION NUMBER: 60/065311		
PRIOR FILING DATE: 1997-11-13	PRIOR FILING DATE: 1997-11-13		
PRIOR APPLICATION NUMBER: 60/066120	PRIOR APPLICATION NUMBER: 60/066120		
PRIOR FILING DATE: 1997-11-21	PRIOR FILING DATE: 1997-11-21		
PRIOR APPLICATION NUMBER: 60/066466	PRIOR APPLICATION NUMBER: 60/066466		
PRIOR FILING DATE: 1997-11-24	PRIOR FILING DATE: 1997-11-24		
PRIOR APPLICATION NUMBER: 60/066772	PRIOR APPLICATION NUMBER: 60/066772		
PRIOR FILING DATE: 1997-11-24	PRIOR FILING DATE: 1997-11-24		
PRIOR APPLICATION NUMBER: 60/069335	PRIOR APPLICATION NUMBER: 60/069335		
PRIOR FILING DATE: 1997-12-11	PRIOR FILING DATE: 1997-12-11		
PRIOR APPLICATION NUMBER: 60/069425	PRIOR APPLICATION NUMBER: 60/069425		
PRIOR FILING DATE: 1997-12-12	PRIOR FILING DATE: 1997-12-12		
PRIOR APPLICATION NUMBER: 60/069870	PRIOR APPLICATION NUMBER: 60/069870		
PRIOR FILING DATE: 1997-12-17	PRIOR FILING DATE: 1997-12-17		
PRIOR APPLICATION NUMBER: 60/068017	PRIOR APPLICATION NUMBER: 60/068017		
PRIOR FILING DATE: 1997-12-18	PRIOR FILING DATE: 1997-12-18		
PRIOR APPLICATION NUMBER: 60/077450	PRIOR APPLICATION NUMBER: 60/077450		
PRIOR FILING DATE: 1998-03-10	PRIOR FILING DATE: 1998-03-10		
PRIOR APPLICATION NUMBER: 60/077632	PRIOR APPLICATION NUMBER: 60/077632		
PRIOR FILING DATE: 1998-03-11	PRIOR FILING DATE: 1998-03-11		
PRIOR APPLICATION NUMBER: 60/077649	PRIOR APPLICATION NUMBER: 60/077649		
PRIOR FILING DATE: 1998-03-11	PRIOR FILING DATE: 1998-03-11		
PRIOR APPLICATION NUMBER: 60/078886	PRIOR APPLICATION NUMBER: 60/078886		
PRIOR FILING DATE: 1998-03-20	PRIOR FILING DATE: 1998-03-20		
PRIOR APPLICATION NUMBER: 60/078939	PRIOR APPLICATION NUMBER: 60/078939		
PRIOR FILING DATE: 1998-03-20	PRIOR FILING DATE: 1998-03-20		
PRIOR APPLICATION NUMBER: 60/079664	PRIOR APPLICATION NUMBER: 60/079664		
PRIOR FILING DATE: 1998-03-27	PRIOR FILING DATE: 1998-03-27		
PRIOR APPLICATION NUMBER: 60/079786	PRIOR APPLICATION NUMBER: 60/079786		
PRIOR FILING DATE: 1998-03-27	PRIOR FILING DATE: 1998-03-27		
PRIOR APPLICATION NUMBER: 60/080107	PRIOR APPLICATION NUMBER: 60/080107		
PRIOR FILING DATE: 1998-03-31	PRIOR FILING DATE: 1998-03-31		
PRIOR APPLICATION NUMBER: 60/080194	PRIOR APPLICATION NUMBER: 60/080194		
PRIOR FILING DATE: 1998-03-31	PRIOR FILING DATE: 1998-03		







GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run On: December 20, 2002, 15:14:53 ; Search time 14.9565 Seconds  
(without alignments)  
519.349 Million cell updates/sec

Title: US-09-856-319B-4  
Perfect score: 1418  
Sequence: 1 MLLLSLTSLVLLSSWGCG.....AMTRYSKFSTWYNQVAYN 264

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	755.5	53.3	231	2	US-09-027-337-6
2	755.5	53.3	231	4	US-09-644-600-6
3	750.5	52.9	230	4	US-08-944-483-62
4	737	52.0	229	2	US-08-557-146-13
5	737	52.0	229	2	US-09-154-344-13
6	724.5	51.1	228	1	US-08-278-091-10
7	724.5	51.1	228	1	US-08-483-859-10
8	724.5	51.1	228	1	US-08-472-173-10
9	724.5	51.1	228	2	US-08-487-167-10
10	724.5	51.1	228	2	US-08-482-816-10
11	724.5	51.1	228	2	US-08-296-149-10
12	724.5	51.1	228	2	US-08-801-499-10
13	724.5	51.1	228	2	US-08-615-271-10
14	724.5	51.1	228	3	US-09-074-660-10
15	724.5	51.1	228	3	US-09-074-659-10
16	724.5	51.1	228	3	US-09-106-468-10
17	724.5	51.1	228	4	US-09-106-466A-10
18	724.5	51.1	228	4	US-09-106-467-10
19	489	34.5	814	1	US-08-750-711-1
20	486	34.3	241	4	US-08-944-483-59
21	485.5	34.2	812	1	US-08-248-629A-1
22	485.5	34.2	812	1	US-08-451-932-1
23	485.5	34.2	812	1	US-08-452-260-1
24	485.5	34.2	812	1	US-08-326-785-1
25	485.5	34.2	812	2	US-08-612-788-1
26	485.5	34.2	812	2	US-08-605-598B-1
27	485.5	34.2	812	2	US-08-429-743-1

28	485.5	34.2	812	2	US-08-866-735-1	Sequence 1, Appli
29	485.5	34.2	812	3	US-09-066-028-1	Sequence 1, Appli
30	485.5	34.2	812	5	PCT-US95-05107-1	Sequence 1, Appli
31	485	34.2	791	1	US-08-643-219-1	Sequence 1, Appli
32	485	34.2	791	3	US-08-851-350-1	Sequence 1, Appli
33	484	34.1	790	6	5200340-6	Patent No. 5200340
34	484	34.1	790	1	US-08-469-486-54	Sequence 54, Appl
35	484	34.1	790	2	US-08-469-658-54	Sequence 54, Appl
36	484	34.1	791	2	US-09-131-995-1	Sequence 1, Appli
37	484	34.1	791	2	US-08-832-087B-1	Sequence 1, Appli
38	484	34.1	791	4	US-09-132-154-1	Sequence 1, Appli
39	484	34.1	810	1	US-07-854-603-2	Sequence 1, Appli
40	484	34.1	810	1	US-08-147-000B-29	Sequence 29, Appl
41	484	34.1	810	4	US-09-086-514-1	Sequence 1, Appli
42	484	34.1	810	6	5200340-8	Patent No. 5200340
43	483.5	34.1	268	1	US-08-568-031-2	Sequence 2, Appli
44	483.5	34.1	268	2	US-08-966-319-2	Sequence 2, Appli
45	483.5	34.1	268	3	US-09-153-304-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-09-027-337-6  
; Sequence 6, Application US/09027337B  
; Patent No. 5972616  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Tanimoto, Hirotooshi  
; TITLE OF INVENTION: TAGD-15: An Extracellular Serine Protease Overexpressed in  
; TITLE OF INVENTION: Breast and Ovarian Carcinomas  
; FILE REFERENCE: D6064  
; CURRENT APPLICATION NUMBER: US/09/027.337B  
; CURRENT FILING DATE: 1998-02-20  
; NUMBER OF SEQ ID NOS: 13  
; SEQ ID NO 6  
; LENGTH: 231  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Serine protease catalytic domain of chymotrypsin (Chymb)  
; OTHER INFORMATION: homologous to similar domain in TAGD-15  
US-09-027-337-6

Query Match	53.3%	Score	755.5	DB	2	Length	231
Best Local Similarity	58.2%	Pred. No.	5.7e-64				
Matches	135	Conservative	35	Mismatches	61	Indels	1
Gaps	1						
Qy	33	RIVGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVTAHCOVTPGRHVVLEGYDR	92				
Db	1	RIVNGEDAVPGSWPQVSLQDKTGFHFCGSLISEDWVYVTAHCGVRTS-DVVVAGEFDQ	59				
Qy	93	SSNAEPVQVLSSTARAITHPNNNANTMNDLTLLKLLASPARVTAQVSPVCLASTNEALPSG	152				
Db	60	GSDEENIQVLKIAKFKPKPSILTVNNDITLLKATAPFSQTVSAVCLPSADDDFFAG	119				
Qy	153	LTCVVTGNGRISGVGNVTPARLQVVLPLVTVQNCQYWGARITDAMICAGSGASSCOG	212				
Db	120	TLCATTGWGKTKYNANKTPDKLQQAALPLLSNAECKSKWGRITDVMICAGASGVSSCMG	179				
Qy	213	DSGGPLVCQKNGTWNLLIGVSGTKNCNIQAPAMTRYSKFSTWYNQVAYN	264				
Db	180	DSGGPLVCQKDGAMTLVGIVSWGSDTCTSSPGVYARVTKLIPWQKILAAAN	231				

RESULT 2  
US-09-644-600-6  
; Sequence 6, Application US/09644600  
; Patent No. 6451500  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Tanimoto, Hirotooshi

;; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease  
;; FILE REFERENCE: Overexpressed in Carcinomas  
;; CURRENT APPLICATION NUMBER: US/09/644,600  
;; CURRENT FILING DATE: 2000-08-23  
;; PRIOR APPLICATION NUMBER: 09/421,213  
;; PRIOR FILING DATE: 1999-10-20  
;; PRIOR APPLICATION NUMBER: 09/027,337  
;; PRIOR FILING DATE: 1998-02-20  
;; NUMBER OF SEQ ID NOS: 98  
;; SEQ ID NO 6  
;; LENGTH: 231  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; OTHER INFORMATION: Chymotrypsin  
US-09-644-600-6

Query Match 53.3%; Score 755.5; DB 4; Length 231;  
Best Local Similarity 58.2%; Pred. No. 5.7e-64;  
Matches 135; Conservative 35; Mismatches 61; Indels 1; Gaps 1;  
  
QY 33 RIVNGENAVPGSPWQVSLQDNTGFHFCGSLISPNWVTAHCVTPGRHFVVLGEYDR 92  
Db 1 RIVNGEDAVPGSPWQVSLQDNTGFHFCGSLISEDMVTAHCGVRTS-DVVYAGEFDQ 59  
  
QY 93 SSNAEPQVLSIARATHPNNWNTMNDTLTKLASPARYTAQVSPVCLASTNEALPSG 152  
Db 60 GSDEENIQVLTKAKVFNKFSLLTVNNDITLLKLATPARFSQTVSAVCLPSADDDFPAG 119  
  
QY 153 LTCVTTGWRISGVNTPARLQOVVPLVTVNQCROYGARITDAMICAGSGSSCOG 212  
Db 120 TLCATTGWGKTKYNAKTPDKLQQAALPLLSNAECKKSGRRITDVMICAGSGVSSCMG 179  
  
QY 213 DSGPLVLCQKNTWLVIGVSGTKNCNIQAAPMYTRVSKFSTWVNOVMAYN 264  
Db 180 DSGPLVLCQKNTWLVIGVSGTKNCNITSSPGVYARVTKLIPWQKILAA 231

RESULT 3  
US-08-944-483-62  
; Sequence 62, Application US/08944483  
; Patent No. 6232456  
; GENERAL INFORMATION:  
; APPLICANT: COHEN, MAURICE  
; APPLICANT: COLPITTS, TRACEY L.  
; APPLICANT: FRIEDMAN, PAULA N.  
; APPLICANT: GRANADOS, EDWARD N.  
; APPLICANT: KLASS, MICHAEL R.  
; APPLICANT: RUSSELL, JOHN C.  
; APPLICANT: STEWART, KENT D.  
; APPLICANT: STROUPE, STEVEN D.  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS  
; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES  
; TITLE OF INVENTION: OF THE PROSTATE  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Abbott Laboratories  
; STREET: 100 Abbott Park Road  
; CITY: Abbott Park  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60064-3500  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/944,483  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Becker, Cheryl L.  
;; REGISTRATION NUMBER: 35,441  
;; REFERENCE/DOCKET NUMBER: 6183.US.01  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 847/935-1729  
;; TELEFAX: 847/938-3623  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 62:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 230 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: No. 6232456e  
US-08-944-483-62  
  
Query Match 52.9%; Score 750.5; DB 4; Length 230;  
Best Local Similarity 58.0%; Pred. No. 1.7e-63;  
Matches 134; Conservative 35; Mismatches 61; Indels 1; Gaps 1;  
  
QY 34 IYNGENAVPGSPWQVSLQDNTGFHFCGSLISPNWVTAHCVTPGRHFVVLGEYDRS 93  
Db 1 IYNGEDAVPGSPWQVSLQDNTGFHFCGSLISEDMVTAHCGVRTS-DVVYAGEFDQ 59  
  
QY 94 SSNAEPQVLSIARATHPNNWNTMNDTLTKLASPARYTAQVSPVCLASTNEALPSGL 153  
Db 60 SDEENIQVLTKAKVFNKFSLLTVNNDITLLKLATPARFSQTVSAVCLPSADDDFPAGT 119  
  
QY 154 TCVTTGWRISGVNTPARLQOVVPLVTVNQCROYGARITDAMICAGSGSSCOGD 213  
Db 120 LCATTGWGKTKYNAKTPDKLQQAALPLLSNAECKKSGRRITDVMICAGSGVSSCMGD 179  
  
QY 214 SGGPLVLCQKNTWLVIGVSGTKNCNIQAAPMYTRVSKFSTWVNOVMAYN 264  
Db 180 SGGPLVLCQKNTWLVIGVSGTKNCNITSSPGVYARVTKLIPWQKILAA 230

RESULT 4  
US-08-557-146-13  
; Sequence 13, Application US/08557146  
; Patent No. 5834290  
; GENERAL INFORMATION:  
; APPLICANT: Egelrud, Torbjorn  
; APPLICANT: Hansson, Lennart  
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic  
; TITLE OF INVENTION: Enzyme (SCCE)  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: White & Case, Patent Department  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2787  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/557,146  
; FILING DATE: 14-DEC-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sterner, Richard J.  
; REGISTRATION NUMBER: 35,372  
; REFERENCE/DOCKET NUMBER: 1103326-181  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 819-8783  
; TELEFAX: (212) 354-8113

INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 229 amino acids  
TYPE: amino acids  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: polypeptide  
US-09-857-146-13

Query Match 52.0%; Score 737; DB 2; Length 229;  
Best Local Similarity 57.6%; Pred. No. 3.2e-62;  
Matches 133; Conservative 36; Mismatches 60; Indels 2; Gaps 2;  
Qy 34 IVNGENAVPGSWPQVSLQDNTGTFHFCGSLISPNWVYTAARHCQVTPGRHFVVLGEYDRS 93  
Db 1 IVNGEDAVPGSWPQVSLQDNTGTFHFCGSLISPNWVYTAARHCQVTPGRHFVVLGEYDRS 59  
Qy 94 SNAEPQVLSIARAITHPNWNTMNDLTLKLLASPARYTAQVSPVCLASTNEALPSGL 153  
Db 60 SDEENIQVLKIAKVEKNPKFSILTVNNDITLLKLATPARFSQTVSAVCLPSADDDFPAGT 119  
Qy 154 TCVTTGGRISGVNVPARLQOVVPLVTVNQCROYNGARITDAMICAGSGASSCGD 213  
Db 120 LCATTGKGKTKYNANKTPDKLQQAALPLLSNAECKSKNGRRITDVMICA-GAGVSSCMGD 178  
Qy 214 SGGPLVCQKGNWTLGIVSGTKNCNIQAPAMYTRVSKSTWNOVMAYN 264  
Db 179 SGGPLVCQKGNWTLGIVSGTKNCNIQAPAMYTRVSKSTWNOVMAYN 229

RESULT 5  
US-09-154-344-13  
Sequence 13, Application US/09154344  
Patent No. 5981256  
GENERAL INFORMATION:  
APPLICANT: Egelrud, Torbjorn  
APPLICANT: Hansson, Lennart  
TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: White & Case, Patent Department  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/154,344  
FILING DATE: 16-SEP-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/557,146  
FILING DATE: 14-DEC-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Sturner, Richard J.  
REGISTRATION NUMBER: 35,372  
REFERENCE/DOCKET NUMBER: 1103326-181  
TELEPHONE: (212) 819-8783  
TELEFAX: (212) 354-8113  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 229 amino acids  
TYPE: amino acids  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: polypeptide  
US-09-154-344-13

Query Match 52.0%; Score 737; DB 2; Length 229;  
Best Local Similarity 57.6%; Pred. No. 3.2e-62;  
Matches 133; Conservative 36; Mismatches 60; Indels 2; Gaps 2;  
Qy 34 IVNGENAVPGSWPQVSLQDNTGTFHFCGSLISPNWVYTAARHCQVTPGRHFVVLGEYDRS 93  
Db 1 IVNGEDAVPGSWPQVSLQDNTGTFHFCGSLISPNWVYTAARHCQVTPGRHFVVLGEYDRS 59  
Qy 94 SNAEPQVLSIARAITHPNWNTMNDLTLKLLASPARYTAQVSPVCLASTNEALPSGL 153  
Db 60 SDEENIQVLKIAKVEKNPKFSILTVNNDITLLKLATPARFSQTVSAVCLPSADDDFPAGT 119  
Qy 154 TCVTTGGRISGVNVPARLQOVVPLVTVNQCROYNGARITDAMICAGSGASSCGD 213  
Db 120 LCATTGKGKTKYNANKTPDKLQQAALPLLSNAECKSKNGRRITDVMICA-GAGVSSCMGD 178  
Qy 214 SGGPLVCQKGNWTLGIVSGTKNCNIQAPAMYTRVSKSTWNOVMAYN 264  
Db 179 SGGPLVCQKGNWTLGIVSGTKNCNIQAPAMYTRVSKSTWNOVMAYN 229

RESULT 6  
US-08-278-091-10  
Sequence 10, Application US/08278091  
Patent No. 5506139  
GENERAL INFORMATION:  
APPLICANT: LOOSMORE, Sheena M  
APPLICANT: YANG, Yan-ping  
APPLICANT: CHONG, Pele  
APPLICANT: OOMEN, Raymond P.  
APPLICANT: KLEIN, Michel H.  
TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with  
NUMBER OF SEQUENCES: 23  
TITLE OF INVENTION: Reduced Protease Activity  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/278,091  
FILING DATE: 21-JUL-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-371  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 228 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-278-091-10

Query Match 51.1%; Score 724.5; DB 1; Length 228;  
Best Local Similarity 57.1%; Pred. No. 4.9e-61;  
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;  
Qy 34 IVNGENAVPGSWPQVSLQDNTGTFHFCGSLISPNWVYTAARHCQVTPGRHFVVLGEYDRS 93

Db 1 IVNGEAVPGSWPQVSLQDKTGFHFCGSLINENWVVTAAHCGVTTSDVVVAGEFDQG 59  
QY 94 SNAEPVQVLSIARAIHPNNWNTMNDLTLKLASPARYTAQVSPVCLASTNEALPSGL 153  
Db 60 SSEKIQKLIKAKVFNKSYNSLTINNDTLKLSTAAFSQTVSAVCLPSASDDFAAGT 119  
QY 154 TCVTTCGWRISGVNVTTPARLQOVLPVTVNOCROYGARITDAMICAGGSSCGD 213  
Db 120 TCVTTCGWLTRYAN-TPDRLOQASLPLLSNTNCKKYWGTKIKDAMICAGAGSVSSCMGD 177  
QY 214 SGGPLVCQKGNWTLGIVSGTKNCNIQAPAMYTRVSKFSTWINOVMAYN 264  
Db 178 SGGPLVCKKNGAWTLGIVSGWSSSTCSTPGVIARVATLVNVMVQOTLAAN 228

## RESULT 7

US-08-483-859-10  
; Sequence 10, Application US/08483859  
; Patent No. 5656436  
; GENERAL INFORMATION:  
; APPLICANT: LOOSMORE, Sheena M.  
; APPLICANT: YANG, Yan-Ping  
; APPLICANT: CHONG, Pele  
; APPLICANT: OOMEN, Raymond P.  
; APPLICANT: KLEIN, Michel H.  
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with  
; TITLE OF INVENTION: Reduced Protease Activity  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: Suite 701, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/483,859  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/296,149  
; FILING DATE: 26-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/278,091  
; FILING DATE: 21-JUL-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I.  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-495 MIS:vg  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 228 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-483-859-10

Query Match 51.1%; Score 724.5; DB 1; Length 228;  
Best Local Similarity 57.1%; Pred. No. 4.9e-61;  
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;  
QY 34 IVNGENAVPGSWPQVSLQDKTGFHFCGSLINENWVVTAAHCGVTTSDVVVAGEFDQS 93  
Db 1 IVNGEAVPGSWPQVSLQDKTGFHFCGSLINENWVVTAAHCGVTTSDVVVAGEFDQG 59

QY 94 SNAEPVQVLSIARAIHPNNWNTMNDLTLKLASPARYTAQVSPVCLASTNEALPSGL 153  
Db 60 SSEKIQKLIKAKVFNKSYNSLTINNDTLKLSTAAFSQTVSAVCLPSASDDFAAGT 119  
QY 154 TCVTTCGWRISGVNVTTPARLQOVLPVTVNOCROYGARITDAMICAGGSSCGD 213  
Db 120 TCVTTCGWLTRYAN-TPDRLOQASLPLLSNTNCKKYWGTKIKDAMICAGAGSVSSCMGD 177  
QY 214 SGGPLVCQKGNWTLGIVSGTKNCNIQAPAMYTRVSKFSTWINOVMAYN 264  
Db 178 SGGPLVCKKNGAWTLGIVSGWSSSTCSTPGVIARVATLVNVMVQOTLAAN 228

## RESULT 8

US-08-472-173-10  
; Sequence 10, Application US/08472173  
; Patent No. 5665353  
; GENERAL INFORMATION:  
; APPLICANT: LOOSMORE, Sheena M.  
; APPLICANT: YANG, Yan-Ping  
; APPLICANT: CHONG, Pele  
; APPLICANT: OOMEN, Raymond P.  
; APPLICANT: KLEIN, Michel H.  
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with  
; TITLE OF INVENTION: Reduced Protease Activity  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: Suite 701, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/472,173  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/296,149  
; FILING DATE: 26-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/278,091  
; FILING DATE: 21-JUL-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I.  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-493 MIS:vg  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 228 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-472-173-10

Query Match 51.1%; Score 724.5; DB 1; Length 228;  
Best Local Similarity 57.1%; Pred. No. 4.9e-61;  
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;  
QY 34 IVNGENAVPGSWPQVSLQDKTGFHFCGSLINENWVVTAAHCGVTTSDVVVAGEFDQS 93  
Db 1 IVNGEAVPGSWPQVSLQDKTGFHFCGSLINENWVVTAAHCGVTTSDVVVAGEFDQG 59  
QY 94 SNAEPVQVLSIARAIHPNNWNTMNDLTLKLASPARYTAQVSPVCLASTNEALPSGL 153

Db 60 SSSEIKQLKIAKFKNSKYNLTINNDITLLKLTAAASFQTSVAVCLPSASDDFAAGT 119  
QY 154 TCVTTGWRISGVGNVTPARLQOVVPLVTVNQCRQYWGARTTDAMICAGSGSSCGD 213  
Db 120 TCVTTGWR-LFRYAN-TPDRLOQASPLLSNTNCKKYGWTKIKDAMICAGASGVSSCMGD 177  
QY 214 SGGPLVCOKGNTWVLIGVSGTKNCNQAPAMYTRVSKFSTWVNOVMAYN 264  
Db 178 SGGPLVCKKNGAWTLVIGVSGSSCTSTPGVYARVTVLNVWVQOTLAAN 228

RESULT 9  
US-08-487-167-10  
: Sequence 10, Application US/08487167  
: Patent No. 5869302  
: GENERAL INFORMATION:  
: APPLICANT: LOOSMORE, Sheena M.  
: APPLICANT: YANG, Yan-Ping  
: APPLICANT: CHONG, Pele  
: APPLICANT: OOMEN, Raymond P.  
: APPLICANT: KLEIN, Michel H.  
: TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with  
: TITLE OF INVENTION: Reduced Protease Activity  
: NUMBER OF SEQUENCES: 23  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Sim & McBurney  
: STREET: Suite 701, 330 University Avenue  
: CITY: Toronto  
: STATE: Ontario  
: COUNTRY: Canada  
: ZIP: M5G 1R7  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/487.167  
: FILING DATE: 07-JUN-1995  
: CLASSIFICATION: 435  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/296.149  
: FILING DATE: 26-AUG-1994  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/278.091  
: FILING DATE: 21-JUL-1994  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Stewart, Michael I.  
: REGISTRATION NUMBER: 24,973  
: REFERENCE/DOCKET NUMBER: 1038-508 MIS:vg  
: TELEPHONE: (416) 595-1155  
: TELEFAX: (416) 595-1163  
: INFORMATION FOR SEQ ID NO: 10:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 228 amino acids  
: TYPE: amino acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
US-08-487-167-10

Query Match 51.1%; Score 724.5; DB 2; Length 228;  
Best Local Similarity 57.1%; Pred. No. 4.9e-61;  
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;  
QY 34 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVYTAACHQVTPGRHFVVLGEYDRS 93  
Db 1 IVNGEAVPGSWPQVSLQDNTGFHFCGSLINWVYTAACHQVTPGRHFVVLGEYDRS 59  
QY 94 SNAEPVQLSTARATHPNWNTNNDITLLKLTAAASFQTSVAVCLPSASDDFAAGT 119  
Db 60 SSSEIKQLKIAKFKNSKYNLTINNDITLLKLTAAASFQTSVAVCLPSASDDFAAGT 119

QY 154 TCVTTGWRISGVGNVTPARLQOVVPLVTVNQCRQYWGARTTDAMICAGSGSSCGD 213  
Db 120 TCVTTGWR-LFRYAN-TPDRLOQASPLLSNTNCKKYGWTKIKDAMICAGASGVSSCMGD 177  
QY 214 SGGPLVCOKGNTWVLIGVSGTKNCNQAPAMYTRVSKFSTWVNOVMAYN 264  
Db 178 SGGPLVCKKNGAWTLVIGVSGSSCTSTPGVYARVTVLNVWVQOTLAAN 228  
RESULT 10  
US-08-482-816-10  
: Sequence 10, Application US/08482816  
: Patent No. 5935573  
: GENERAL INFORMATION:  
: APPLICANT: LOOSMORE, Sheena M  
: APPLICANT: YANG, Yan-Ping  
: APPLICANT: CHONG, Pele  
: APPLICANT: OOMEN, Raymond P.  
: APPLICANT: KLEIN, Michel H.  
: TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with Reduced Protease  
: NUMBER OF SEQUENCES: 23  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Sim & McBurney  
: STREET: Suite 701, 330 University Avenue  
: CITY: Toronto  
: STATE: Ontario  
: COUNTRY: Canada  
: ZIP: M5G 1R7  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/482.816  
: FILING DATE: 07-JUN-1995  
: CLASSIFICATION: 435  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/296.149  
: FILING DATE: 26-AUG-1994  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/278.091  
: FILING DATE: 21-JUL-1994  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Stewart, Michael I.  
: REGISTRATION NUMBER: 24,973  
: REFERENCE/DOCKET NUMBER: 1038-494 MIS:vg  
: TELEPHONE: (416) 595-1155  
: TELEFAX: (416) 595-1163  
: INFORMATION FOR SEQ ID NO: 10:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 228 amino acids  
: TYPE: amino acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
US-08-482-816-10

Query Match 51.1%; Score 724.5; DB 2; Length 228;  
Best Local Similarity 57.1%; Pred. No. 4.9e-61;  
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;  
QY 34 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVYTAACHQVTPGRHFVVLGEYDRS 93  
Db 1 IVNGEAVPGSWPQVSLQDNTGFHFCGSLINWVYTAACHQVTPGRHFVVLGEYDRS 59  
QY 94 SNAEPVQLSTARATHPNWNTNNDITLLKLTAAASFQTSVAVCLPSASDDFAAGT 119  
Db 60 SSSEIKQLKIAKFKNSKYNLTINNDITLLKLTAAASFQTSVAVCLPSASDDFAAGT 119  
QY 154 TCVTTGWRISGVGNVTPARLQOVVPLVTVNQCRQYWGARTTDAMICAGSGSSCGD 213

RECEIVED 11

APPLICANT: COMEN, Raymond F.  
 APPLICANT: KLEIN, Michel H.  
 TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with Reduced Protease A  
 NUMBER OF SEQUENCES: 23

```

;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,499
; FILING DATE:
;

```

```

/ APPLICATION NUMBER: 08-08482,816
/ PRIORITY DATE: 07-JUN-1995
/ PRIORITY APPLICATION DATA: 08/296,149
/ APPLICATION NUMBER: 08-08482,816
/ PRIORITY DATE: 07-JUN-1995
/ PRIORITY APPLICATION DATA: 08/296,149

```

: FILING DATE: 26-AUG-1994  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 08/278,091

FILING DATE: 21-JUL-1994

ATTORNEY/AGENT INFORMATION:

```

; NAME: Stewart, Michael I

```

; REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-6711

; TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163

; INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 228 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLGI: Linear  
;  
PC-08-807-499-10

03 00 001 433 10

Query Match

Best Local Similarity 57.1%: Pred. N 5000

Matches 132: Conservative 34: Mism

.....

QY : 34 IVGENAVPGSWPQVSLQDNTGFHFCCGSL

100

Ddb 1 IVNGEEAVPGSWPQVSLQDKTGFHFCCGSL

Qy 94 SNAEPVQLSIARATHPNWNANTMNNDLTL

[illegible]

Db 60 SSSEKIQKLKIAKVFKNSKYNLSLTINNDITL

QY 154 TCVTIGWGRISGVGNVT<sup>1</sup>PARLQQWLP<sup>1</sup>LVTV

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

DB I20 TCVITGWG-LIRYAN-TIPDRLQQASLP LLSN

[illegible]

QY ZI4 SGGPLVCQAGNTWLTGIVSWGIAKNCNIQAF

178 S C C R I V C K K N C A W T T I V C T V S W G C C T C C T S T E P B

DD T/S DDFFACGKNNYNYWHTIMDATIDTOMCAIICISITIE

## RESULT 13

1000

```

US-08-615-271-10
; Sequence 10, Application US/08615271
; Patent No. 5981503
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED
; TITLE OF INVENTION: PROTEASE ACTIVITY
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/615,271
; FILING DATE: 20-JUN-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-580
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-615-271-10

```

```

Query Match 51.1%; Score 724.5; DB 2; Length 228;
Best Local Similarity 57.1%; Pred. No. 4.9e-61;
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;

Qy 34 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVYTAHRCQVTPGRHFVVLGEYDRS 93
Db 1 IVNGEEAVPGSWPQVSLQDNTGFHFCGSLINENWVYTAHRCVTTTS-DVVVAGEFDQG 59

Qy 94 SNAEPVQLSTARATHPNWNTNNDTLTKLASPARYTAQVSPVCLASTNEALPSGL 153
Db 60 SSSEKIQLKIAKFKNSKYNSTINNDITLLKLTAAASFQTSVAVCLPSASDDFAAGT 119

Qy 154 TCVTTGWRIGSGVNTTPARLQVVLPLVTYNOCROYGARTIDAMICAGSGGSSCCGD 213
Db 120 TCVTTGWR-LTRYAN-TPDRLOQASLPLLSNTNCKYKYGTKIKDAMICAGASGVSSCMGD 177

Qy 214 SGGPLVCKKNGAWTLVGIVSGWSTCSSTSTPGVYARVYALVNWVQOTLAAN 228
Db 178 SGGPLVCKKNGAWTLVGIVSGWSTCSSTSTPGVYARVYALVNWVQOTLAAN 228

RESULT 14
US-09-074-660-10
; Sequence 10, Application US/09074660
; Patent No. 6020183
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OOMEN, Raymond P.

```

```

; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; TITLE OF INVENTION: Reduced Protease Activity
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/074,660
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,167
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,149
; FILING DATE: 26-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,091
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-731 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-074-660-10

Query Match 51.1%; Score 724.5; DB 3; Length 228;
Best Local Similarity 57.1%; Pred. No. 4.9e-61;
Matches 132; Conservative 34; Mismatches 62; Indels 3; Gaps 3;

Qy 34 IVNGENAVPGSWPQVSLQDNTGFHFCGSLISPNWVYTAHRCQVTPGRHFVVLGEYDRS 93
Db 1 IVNGEEAVPGSWPQVSLQDNTGFHFCGSLINENWVYTAHRCVTTTS-DVVVAGEFDQG 59

Qy 94 SNAEPVQLSTARATHPNWNTNNDTLTKLASPARYTAQVSPVCLASTNEALPSGL 153
Db 60 SSSEKIQLKIAKFKNSKYNSTINNDITLLKLTAAASFQTSVAVCLPSASDDFAAGT 119

Qy 154 TCVTTGWRIGSGVNTTPARLQVVLPLVTYNOCROYGARTIDAMICAGSGGSSCCGD 213
Db 120 TCVTTGWR-LTRYAN-TPDRLOQASLPLLSNTNCKYKYGTKIKDAMICAGASGVSSCMGD 177

Qy 214 SGGPLVCKKNGAWTLVGIVSGWSTCSSTSTPGVYARVYALVNWVQOTLAAN 228
Db 178 SGGPLVCKKNGAWTLVGIVSGWSTCSSTSTPGVYARVYALVNWVQOTLAAN 228

RESULT 15
US-09-074-659-10
; Sequence 10, Application US/09074659
; Patent No. 6025342
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele

```

Search completed: December 20, 2002, 15:20:37  
Job time : 15.9565 secs



GenCore version 5.1.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:14:08 ; Search time 32.6957 Seconds  
(without alignments)  
1663.721 Million cell updates/sec

Title: US-09-856-319b-4  
Perfect score: 1418  
Sequence: 1 MLLLSLTLVLGSSWGCG.....AMVTRVSKFSTWVQVMAYN 264

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriaph.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1418	100.0	264	11 Q9ER05	Q9er05 mus musculus
2	1414	99.7	264	11 Q9D7P8	Q9d7p8 mus musculus
3	1408	99.3	264	11 Q9D960	Q9d960 mus musculus
4	1365	96.3	264	11 Q9EQ28	Q9eq28 rattus norv
5	894.5	63.1	261	13 Q9W7Q4	Q9w7q4 paralichthy
6	805.5	56.8	260	13 Q9W7Q3	Q9w7q3 paralichthy
7	797.5	56.2	263	11 Q9DC86	Q9dc86 mus musculus
8	795.5	56.1	263	11 Q9CR35	Q9cr35 mus musculus
9	789.5	55.7	263	11 Q9D8X8	Q9d8x8 mus musculus
10	758.5	53.5	263	13 Q9PWQ6	Q9pwq6 gadus morhu
11	674	47.5	164	11 Q9DC82	Q9dc82 mus musculus
12	552.5	39.0	269	11 Q9CQ52	Q9cq52 mus musculus
13	544.5	38.4	269	11 Q9D7T9	Q9d7t9 mus musculus
14	533.5	37.6	266	13 Q9Z077	Q9z077 gadus morhu
15	524.5	37.0	260	13 Q9W7P9	Q9w7p9 paralichthy
16	510	36.0	267	5 Q9BK47	Q9bk47 luidia foli

# ALIGNMENTS

RESULT 1

ID	Q9ER05	PRELIMINARY;	PRT;	264 AA.
AC	Q9ER05;			
DT	01-MAR-2001 (Tremblrel. 16, Created)			
DT	01-MAR-2001 (Tremblrel. 16, Last sequence update)			
DT	01-JUN-2002 (Tremblrel. 21, Last annotation update)			
DE	Chymopasin (Chymotrypsin A CTRA-1).			
GN	CTRL OR CTRL1.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Mitsui S., Yamaguchi N.;			
RT	"Molecular cloning of mouse chymopasin.";			
RL	Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=129S6/SVEVTAC; TISSUE=SPLEEN;			
RA	Bjoernslett M.;			
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases			
CC	- - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.			
DR	EMBL; AB016228; BAB20275.1; -			
DR	EMBL; AF236365; AAL11034.1; -			
DR	HSSP; P00766; 4CHA.			
DR	MEROPS; S01.256; -			
DR	MGD; MGI:88558; Ctrl.			
DR	InterPro; IPR001314; Chymotrypsin.			
DR	InterPro; IPR001254; Ser_protease_try.			
DR	Pfam; PF00089; trypsin; 1.			
DR	PRINTS; PR00722; CHYMOTRYPSIN.			
DR	SMART; SM00020; Tryp_SPC; 1.			
DR	PROSITE; PS00240; TRYPSIN_DOM; 1.			
DR	PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.			
DR	PROSITE; PS00135; TRYPSIN_SER; 1.			
KW	Hydrolase; Serine protease.			

Q95kw7 bos taurus  
Q9w7q1 paralichthy  
Q91674 xenopus lae  
Q9w7q0 paralichthy  
Q91039 gadus morhu  
Q91wj5 mus musculus  
Q15146 homo sapien  
Q96ql8 homo sapien  
Q96qv5 homo sapien  
Q96899 scolopendra  
Q19023 macaca mula  
Q46507 papio hamad  
Q9w7q2 paralichthy  
Q99144 mus musculus  
Q9r0w3 rattus norv  
Q921n4 mus musculus  
Q18783 macropus eu  
Q8qgf6 xenopus lae  
Q8vde0 mus musculus  
Q46506 papio hamad  
Q42608 petromyzon  
Q42158 petromyzon  
Q91x79 mus musculus  
Q46644 macaca fasc  
Q921r9 mus musculus  
Q8rla6 mus musculus  
Q9qk9 mus musculus  
Q8ggw3 anguilla ja  
Q920k3 rattus norv

17	508.5	35.9	269	6	Q95KW7
18	503	35.5	249	13	Q9W7Q1
19	499	35.2	1524	13	Q91674
20	498	35.1	266	13	Q9W7Q0
21	494	34.8	270	13	Q91039
22	488.5	34.4	812	11	Q91WJ5
23	485	34.2	810	4	Q15146
24	484.5	34.2	270	4	Q96QL8
25	483	34.1	269	4	Q96QV5
26	473	33.4	277	5	Q96899
27	470.5	33.2	257	6	Q19023
28	468	33.0	334	6	Q46507
29	467	32.9	268	13	Q9W7Q2
30	466.5	32.9	339	11	Q99144
31	466.5	32.9	812	11	Q9R0W3
32	465.5	32.8	273	11	Q921N4
33	456	32.2	806	6	Q18783
34	455.5	32.1	264	13	Q8QGF6
35	455	32.0	453	11	Q8VDE0
36	454	32.0	454	6	Q46506
37	452.5	31.9	247	13	Q42608
38	451.5	31.8	247	13	Q42158
39	450	31.7	266	11	Q91X79
40	449	31.7	266	6	Q46644
41	446.5	31.5	246	11	Q921R9
42	446.5	31.5	331	11	Q8RLA6
43	445.5	31.4	246	11	Q9QK9
44	443.5	31.3	244	13	Q8GGW3
45	442.5	31.2	490	11	Q920K3

```

SQ SEQUENCE 264 AA; 28135 MW; 1D979709A07056C2 CRC64;
Query Match 100.0%; Score 1418; DB 11; Length 264;
Best Local Similarity 100.0%; Pred. No. 1.4e-118;
Matches 264; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLGSSGCGVPAITPALSYNQRIYNGENAVPGSWPQVSLQDNTGFHFC 60
Db 1 MLLSLTSLVLGSSGCGVPAITPALSYNQRIYNGENAVPGSWPQVSLQDNTGFHFC 60
QY 61 GGSLLSPNWWVTAACHCQVTPGRHFVVLGEYDRSSNAEPQVLSIARATHPNWNTMNN 120
Db 61 GGSLLSPNWWVTAACHCQVTPGRHFVVLGEYDRSSNAEPQVLSIARATHPNWNTMNN 120
QY 121 DLTLLKLSPARYTAQVSPVCLASTNEALPSGLTCTVTTGWGRISGVGNVTPARLQOVVLP 180
Db 121 DLTLLKLSPARYTAQVSPVCLASTNEALPSGLTCTVTTGWGRISGVGNVTPARLQOVVLP 180
QY 181 LVTVNQCROYWGARTDAMICAGSGGSCQDGGPLVCQKNTWVLIGIVSWGTKNCN 240
Db 181 LVTVNQCROYWGARTDAMICAGSGGSCQDGGPLVCQKNTWVLIGIVSWGTKNCN 240
QY 241 IQAPAMYTRVSKFSTWVINOVMAYN 264
Db 241 IQAPAMYTRVSKFSTWVINOVMAYN 264

RESULT 2
Q9D7P8 PRELIMINARY; PRT; 264 AA.
AC Q9D7P8;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 1810004D15R1k protein.
GN CTRL OR 1810004D15R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Schriml L.M., Staublei F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sakai H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AK009019; BAB26029.1; -.
DR HSSP; P00766; 4CHA.
DR MEROPS; S01.256; -.
DR MGD; MGI:88558; Ctrl.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
```

```

DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRY_PSP; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 264 AA; 28167 MW; 1D979469A07056C2 CRC64;
Query Match 99.7%; Score 1414; DB 11; Length 264;
Best Local Similarity 99.6%; Pred. No. 3.1e-118;
Matches 263; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLGSSGCGVPAITPALSYNQRIYNGENAVPGSWPQVSLQDNTGFHFC 60
Db 1 MLLSLTSLVLGSSGCGVPAITPALSYNQRIYNGENAVPGSWPQVSLQDNTGFHFC 60
QY 61 GGSLLSPNWWVTAACHCQVTPGRHFVVLGEYDRSSNAEPQVLSIARATHPNWNTMNN 120
Db 61 GGSLLSPNWWVTAACHCQVTPGRHFVVLGEYDRSSNAEPQVLSIARATHPNWNTMNN 120
QY 121 DLTLLKLSPARYTAQVSPVCLASTNEALPSGLTCTVTTGWGRISGVGNVTPARLQOVVLP 180
Db 121 DLTLLKLSPARYTAQVSPVCLASTNEALPSGLTCTVTTGWGRISGVGNVTPARLQOVVLP 180
QY 181 LVTVNQCROYWGARTDAMICAGSGGSCQDGGPLVCQKNTWVLIGIVSWGTKNCN 240
Db 181 LVTVNQCROYWGARTDAMICAGSGGSCQDGGPLVCQKNTWVLIGIVSWGTKNCN 240
QY 241 IQAPAMYTRVSKFSTWVINOVMAYN 264
Db 241 IQAPAMYTRVSKFSTWVINOVMAYN 264

RESULT 3
Q9D960 PRELIMINARY; PRT; 264 AA.
ID Q9D960
AC Q9D960;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 1810004D15R1k protein.
GN CTRL OR 1810004D15R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Schriml L.M., Staublei F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sakai H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AK007333; BAB24967.1; -.
```

```
DR HSSP; P00766; 4CHA.
DR MEROPS; S01.256; -.
DR MGD; MGI-88558; Ctrl.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS02040; TRYP_SIN_DOM; 1.
DR PROSITE; PS00134; TRYP_SIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYP_SIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 264 AA; 28151 MW; 1D979719E07C16DE CRC64;

Query Match          99.3%; Score 1408; DB 11; Length 264;
Best Local Similarity 99.6%; Pred. No. 1.le-117;
Matches 263; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLLSLTSLVLLGSSWCGVPAITPALSYNQRIYNGENAVPGSWPQVSLQDNTGFHFC 60
Db 1 MLLSLTSLVLLGSSWCGVPAITPALSYNQRIYNGENAVPGSWPQVSLQDNTGFHFC 60

Qy 61 GGSLSIPNNVYTAHCQVTPGRHFVVLGEYDRSSNAEPQVLSIARATHPNWNTMNN 120
Db 61 GGSLSIPNNVYTAHCQVTPGRHFVVLGEYDRSSNAEPQVLSIARATHPNWNTMNN 120

Qy 121 DLTLLKASPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQOVVLP 180
Db 121 DLTLLKASPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQOVVLP 180

Qy 181 LVTVNQCROYWGARTIDAMICAGSGASSCGDGGPLVCQKGNWTWVLIGIVSWGTKNCN 240
Db 181 LVTVNQCROYWGARTIDAMICAGSGASSCGDGGPLVCQKGNWTWVLIGIVSWGTKNCN 240

Qy 241 IOAPAMYTRVSKFSTWVNOVMAYN 264
Db 241 IOAPAMYTRVSKFSTWVNOVMAYN 264

RESULT 4
Q9EQZ8 PRELIMINARY; PRT; 264 AA.
ID Q9EQZ8;
AC Q9EQZ8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chymopasin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=RAT PANCREAS;
RA Sogame Y., Mitsui S., Kataoka K., Kashima K., Kato M., Sakagami J.,
RA Yamaguchi N.;
RT "Molecular cloning of rat chymopasin.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYP_SIN FAMILY.
DR EMBL; AB020757; BAB20287.1; -.
DR HSSP; P00766; 4CHA.
DR MEROPS; S01.256; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS02040; TRYP_SIN_DOM; 1.
DR PROSITE; PS00134; TRYP_SIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYP_SIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 264 AA; 28116 MW; F9ED5D210FD3500E CRC64;
```

```
Query Match          96.3%; Score 1365; DB 11; Length 264;
Best Local Similarity 93.6%; Pred. No. 7.2e-114;
Matches 247; Conservative 16; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLLSLTSLVLLGSSWCGVPAITPALSYNQRIYNGENAVPGSWPQVSLQDNTGFHFC 60
Db 1 MLLSLTSLVLLGSSWCGVPAITPALSYNQRIYNGENAVPGSWPQVSLQDNTGFHFC 60

Qy 61 GGSLSIPNNVYTAHCQVTPGRHFVVLGEYDRSSNAEPQVLSIARATHPNWNTMNN 120
Db 61 GGSLSIPNNVYTAHCQVTPGRHFVVLGEYDRSSNAEPQVLSIARATHPNWNTMNN 120

Qy 121 DLTLLKASPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQOVVLP 180
Db 121 DLTLLKASPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQOVVLP 180

Qy 181 LVTVNQCROYWGARTIDAMICAGSGASSCGDGGPLVCQKGNWTWVLIGIVSWGTKNCN 240
Db 181 LVTVNQCROYWGARTIDAMICAGSGASSCGDGGPLVCQKGNWTWVLIGIVSWGTKNCN 240

Qy 241 IOAPAMYTRVSKFSTWVNOVMAYN 264
Db 241 IOAPAMYTRVSKFSTWVNOVMAYN 264

RESULT 5
Q9W7Q4 PRELIMINARY; PRT; 261 AA.
ID Q9W7Q4;
AC Q9W7Q4;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chymotrypsinogen 1.
OS Paralicthys olivaceus (Flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidae; Paralicthyidae; Paralicthys.
OX NCBI_TaxID=8255;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA Suzuki T., Srivastava A.S., Kurokawa T.;
RT "Japanese flounder mRNA for chymotrypsinogen 1.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYP_SIN FAMILY.
DR EMBL; AB029753; BAA82365.1; -.
DR HSSP; P00763; 1DPO.
DR MEROPS; S01.256; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS02040; TRYP_SIN_DOM; 1.
DR PROSITE; PS00134; TRYP_SIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYP_SIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 261 AA; 28184 MW; D7090A9D65395B7D CRC64;

Query Match          63.1%; Score 894.5; DB 13; Length 261;
Best Local Similarity 63.0%; Pred. No. 6.6e-72;
Matches 167; Conservative 33; Mismatches 60; Indels 5; Gaps 3;

Qy 1 MLLSLTSLVLLGSSWCGVPAITPALSYNQRIYNGENAVPGSWPQVSLQDNTGFHFC 60
Db 1 MLLFISCFALV--ASALGCGVPSIKPQVSYGNKYNGETAVSGSWPQVSLQDGRGFHFC 58

Qy 61 GGSLSIPNNVYTAHCQVTPGRHFVVLGEYDRSSNAEPQVLSIARATHPNWNTMNN 120
Db 59 GGSLSIPYVWVTAACHTSPRHRHVLGSDHROYNNEPIQVMSIARATHPYNSQNFNN 118
```



```

Db 180 IVSEAKCKESWGSKITDVMICAGASGVSSCMGDSGGPLVCQKDGWTLGIVSGSGFCS 239
Qy 241 IQAPAMTRVSKFSTWVNOVMAYN 264
Db 240 TSTPAVYARVATLMPWQVQILEAN 263

RESULT 8
Q9CR35 Q9CR35 PRELIMINARY; PRT; 263 AA.
AC Q9CR35;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 2200008D09Rik protein.
GN 2200008D09Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS7BL/6J; TISSUE=STOMACH, SPLEEN, AND PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikola I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPsin FAMILY
DR EMBL; AK008927; BAB25971.1; -
DR EMBL; AK003079; BAB22553.1; -
DR EMBL; AK007765; BAB25241.1; -
DR EMBL; AK007815; BAB25280.1; -
DR EMBL; AK008729; BAB25861.1; -
DR EMBL; AK008888; BAB25954.1; -
DR HSSP; P00766; LGCT.
DR MEROPS; S01.152; -.
DR MGD; MGI:1913723; 2200008D09Rik.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS02040; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 263 AA; 27822 MW; 28C4487AF1A36B27 CRC64;

Query Match 56.18; Score 795.5; DB 11; Length 263;
Best Local Similarity 54.98; Pred. No. 4.5e-63;
Matches 145; Conservative 41; Mismatches 77; Indels 1; Gaps 1;

Qy 1 MLLSLTLSLVLLGSSWCGVPAITPALSYNQRIYNGENAVPGSWPQVSLQDNTGFHFC 60

```

```

Db 1 MAFLLVLSFCALVGATFCGVPATOPVLTLGLSRIVNGEDATPGSWPQVSLQDNTGFHFC 60
Qy 61 GGSLLSPNVTAAHCQVTPGRHFVVLGEYDRSSNAEPQVLSIARATHPNWNNANTWNN 120
Db 61 GGSLLSPNVTAAHCQVTPGRHFVVLGEYDRSSNAEPQVLSIARATHPNWNNANTWNN 119
Qy 121 DLTLLKLASPARYTAQVSPKLASTNEALPSGLCTVTTGWRISGVGNVTARLQOVVLP 180
Db 120 DITLLKLATPAQFSETVSANCLPTVDDFFPAGTLCATGKTKYNAUKTDPKLOOAAALP 179
Qy 181 LVTYNQCRQYWGARITDAMICAGSGASSCGDGGPLVCQKGNVTWLVIGVSWGTKNCN 240
Db 180 IVSEAKCKESWGSKITDVMICAGASGVSSCMGDSGGPLVCQKDGWTLGIVSGSGFCS 239
Qy 241 IQAPAMTRVSKFSTWVNOVMAYN 264
Db 240 TSTPAVYARVATLMPWQVQILEAN 263

```

## RESULT 9

```

Q9D8X8 Q9D8X8 PRELIMINARY; PRT; 263 AA.
AC Q9D8X8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 2200008D09Rik protein.
GN 2200008D09Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS7BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikola I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPsin FAMILY.
DR EMBL; AK007566; BAB25112.1; -
DR HSSP; P00766; LGCT.
DR MEROPS; S01.152; -.
DR MGD; MGI:1913723; 2200008D09Rik.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS02040; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 263 AA; 27898 MW; C0638FB8F905A92F CRC64;

```

Query Match 55.7%; Score 789.5; DB 11; Length 263;

Best Local Similarity 54.5%; Pred. No. 1.6e-62;

Matches 144; Conservative 41; Mismatches 78; Indels 1; Gaps 1;

QY 1 MLLSLTSLVLGSSWGGVPAITPALSYNQRIVNGENAVPGSWPQVSLQDNTGTFHFC 60

DB 1 MAFLLVSCFALVATFGCGVPAIQVLTGLSRVNGEDAIPGSPWQVSLQDRTGFHFC 60

QY 61 GGLISPNWVVAHQCQVTPGRHFVVLGVDYDRSSNAEPQVLSIARAITHPNWNANTMNN 120

DB 61 GGLISENWWVAHCGVKT-TDVVVGEEFDGSDENQVQLKIAOVFNKPFNSFTVRN 119

QY 121 DITLLKLAPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQOVLP 180

DB 120 DITLLKLATPAQSETVSAVCLTVDDEPAGTLCATTCGKTKYNALKTDPKLOQAALP 179

QY 181 LVTVNCQRYGARIIDAMICAGGASSCGSGPLVCQKGNVTWVLIGIVSWGKNCN 240

DB 180 IVSEAKCKESWGSKIITDVNICAGASGVSSCMGDSGGLVCQKQDVWVTLAGIVSWGSGFCS 239

QY 241 IOAPAMVTRVSTWNOVMAYN 264

DB 240 TSTPAYARVATLMPVQELLEAN 263

RESULT 10

Q9PWQ6 PRELIMINARY; PRT; 263 AA.

AC Q9PWQ6

DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-MAY-2002 (TREMBlrel. 20, Last annotation update)

DE Chymotrypsin B precursor (EC 3.4.21.1).

GN CHYB

OS Gadus morhua (Atlantic cod).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.

OX NCBI\_TaxID=8049;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=PYLORIC CARCA;

RA MEDLINE=2046334; PubMed=11011764;

RA Spilliaert R., Gudmundsdottir A.;

RT "Molecular Cloning of the Atlantic Cod Chymotrypsinogen B.;"

RL Microb. Comp. Genomics 5:41-50(2000).

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE

TRYPSIN FAMILY.

CC EMBL: AJ242521; CAB43766.1; .

DR HSSP: P00766; ICHG.

DR MEROPS: S01.152; .

DR InterPro: IPR001314; Chymotrypsin.

DR InterPro: IPR001254; Ser.protease\_Try.

DR Pfam: PF00089; trypsin; 1

DR PRINTS: PR00722; CHYMOTRYPSIN.

DR SMART: SM00020; Tryp.Spc; 1.

DR PROSITE: PS00240; TRYPSIN\_DOM; 1.

DR PROSITE: PS00134; TRYPSIN\_HIS; UNKNOWN\_1.

DR PROSITE: PS00135; TRYPSIN\_SER; 1.

KW Hydrolase; Serine protease; Signal.

FT SIGNAL 1 16 POTENTIAL.

FT CHAIN 17 263 CHYMOTRYPSIN B.

SQ SEQUENCE 263 AA; 28175 MW; EF61B18A34EE5E7C CRC64;

Query Match 53.5%; Score 758.5; DB 13; Length 263;

Best Local Similarity 53.5%; Pred. No. 9.1e-60;

Matches 137; Conservative 47; Mismatches 65; Indels 7; Gaps 5;

QY 13 LGSSWCGVPAITPALSYNQRIVNGENAVPGSWPQVSLQDNTGTFHFCGSGSLISPNWVYT 72

DB 11 ISAAYGCGPAIQPQVTVGARIVNGEAVPHSPWQVSLQSQSGFPHFCGSGSLINENWVYT 70

QY 73 AANCQVTPGRHFVVLGVEYDRES-SNAEPQVLSIARAITHPNWNANTMNDLTLKLASPA 131

DB 71 AANCNVRT-YHRVIVGEHDKSRASDENIQLKPSWFTTHPKWDSRTINNDLSIKLASPA 129

QY 132 RYTAQVSPVCLASTNEALPSGLTCVTTGNG--RISGVGNVTPARLQOVLPVLTVNOCRQ 189

DB 130 VLGTNVSPVCLGESSDVFAFGKCVTSWGMLTRYNAPG--TPNKLQAAALPLMSNECSQ 187

QY 190 YNGAR-ITDAMICAGGASSCGSGPLVCQKGNVTWVLIGIVSWGKNCNIAQAPAMYT 248

DB 188 TWGNMISDMICAGAAGATSCMGDSGGLVCQKDNVTLGVISWSSRCSVTTTAVYA 247

QY 249 RVSKFSTWNOVMAYN 264

DB 248 RVTGLRGWVDQILAN 263

RESULT 11

Q9DC82 PRELIMINARY; PRT; 164 AA.

AC Q9DC82

DT 01-JUN-2001 (TREMBlrel. 17, Created)

DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Q10001G08RIK protein.

GN CTRL OR 0910001G08RIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=SPLEEN;

RA MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustingich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,

RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,

RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.;"

RL Nature 409:685-690(2001).

DR EMBL: AK003074; BAB22549.1; .

DR HSSP: P00763; IDPO.

DR MEROPS: S01.997; .

DR MGD: MGI:88558; Ctrl.

DR InterPro: IPR001254; Ser.protease\_Try.

DR Pfam: PF00089; trypsin; 1.

DR SMART: SM00020; Tryp.Spc; 1.

DR PROSITE: PS50240; TRYPSIN\_DOM; 1.

DR PROSITE: PS00135; TRYPSIN\_SER; 1.

KW Hydrolase; Serine protease.

SQ SEQUENCE 164 AA; 17707 MW; 83791FD829ABEDD6 CRC64;

Query Match 47.5%; Score 674; DB 11; Length 164;

Best Local Similarity 82.2%; Pred. No. 1.7e-52;

Matches 129; Conservative 5; Mismatches 21; Indels 2; Gaps 2;

QY 110 HPNWN-NTMNDLTLKLKASP-ARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGVG 167

DB 8 HPWGGPHLKKTRLFLEFFFPPEKTKTFFPGLASTNEALPSGLTCVTTGWRISGVG 67









Db 175 ATCTRSDWNGSLVTNSMYCAGAGDLASCNGDSGGPLACQSPDGSWEVHGYYVFGSSMGC 234  
QY 240 NI-QAPAMYTRYSKFSTWINOVM 261  
Db 235 NYPKAPSVFTRVSAYIPWINVM 257

Search completed: December 20, 2002, 15:18:54  
Job time : 33.6957 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:12:13 ; Search time 40.6957 Seconds  
(without alignments)  
864.421 Million cell updates/sec

Title: US-09-856-319B-4

Perfect score: 1418

Sequence: 1 MLLLSLTSLVLLGSSWGCG.....AMTRYSKFSTWNOVMAYN 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1418	100.0	264	21	AA11711
2	1262	89.0	264	21	AA11710
3	889	62.7	192	21	AA194077
4	822.5	58.0	263	23	AAU82738
5	774.5	54.6	245	21	AA199596
6	755.5	53.3	231	22	AA198504
7	537	37.9	269	7	AA1981724
8	537	37.9	269	8	AA1981724
9	523.5	36.9	253	13	AA1981724
10	515	36.3	269	7	AA1981724

11	500	35.3	279	22	ABG20513	Novel human diapo
12	498	35.1	269	7	AA1981723	Human elastase II.
13	496.5	35.0	268	17	AA1981723	Human caldesmon co
14	492	34.7	252	7	AA1981723	Sequence of human
15	487.5	34.4	811	12	AA1981723	Plasminogen muten
16	486.5	34.3	268	17	AA1981723	Rat caldesmon cont
17	486.5	34.3	811	12	AA1981723	Plasminogen muten
18	486.5	34.3	811	12	AA1981723	Plasminogen muten
19	485.5	34.2	270	8	AA1981723	Human pancreas ela
20	485.5	34.2	811	12	AA1981723	Plasminogen muten
21	485.5	34.2	812	16	AA1981723	Complete mouse pla
22	485.5	34.2	812	17	AA1981723	Murine plasminogen
23	485.5	34.2	812	20	AA1981723	Murine plasminogen
24	485.5	34.2	812	20	AA1981723	Murine plasminogen
25	485.5	34.2	812	21	AA1981723	Murine plasminogen
26	485.5	34.2	812	21	AA1981723	Murine plasminogen
27	485.5	34.2	812	22	AA1981723	Murine plasminogen
28	485.5	34.2	812	22	AA1981723	Amino acid sequenc
29	485	34.2	791	18	AA1981723	Human plasminogen
30	485	34.2	810	23	AA1981723	Human plasminogen
31	485	34.2	810	23	AA1981723	Human plasminogen
32	485	34.2	1220	22	AA1981723	Novel human secret
33	484.5	34.2	270	23	AA1981723	Human pancreatic t
34	484.5	34.2	273	21	AA1981723	Human pancreatic c
35	484.5	34.2	811	12	AA1981723	Plasminogen muten
36	484	34.1	249	23	AA1981723	Human microplasma
37	484	34.1	348	23	AA1981723	Human miniplasma
38	484	34.1	437	19	AA1981723	Human plasminogen
39	484	34.1	546	14	AA1981723	Sequence of tissue
40	484	34.1	790	15	AA1981723	Human 'Glu' plasm
41	484	34.1	790	22	AA1981723	Mammalian kringl
42	484	34.1	791	21	AA1981723	Human plasminogen
43	484	34.1	791	21	AA1981723	Human plasminogen
44	484	34.1	791	21	AA1981723	Human plasminogen
45	484	34.1	791	22	AA1981723	Amino acid sequenc

ALIGNMENTS

RESULT 1  
AA11711  
ID AA11711 standard; Protein: 264 AA.

AC AA11711;

DT 23-OCT-2000 (first entry)

XX Mouse serine protease BSSP5 (mBSSP5) SEQ ID NO:4.

DE BSSP5; serine protease; human; hBSSP5; mouse; mBSSP5; brain;  
KW diagnostic marker; antibody; transgenic animal; Alzheimer's disease;  
KW epilepsy; cancer; inflammation; infertility; pancreatitis;  
KW prostatic hypertrophy.

OS Mus sp.

PN WO200031243-A1.

XX 02-JUN-2000.

XX 19-NOV-1999; 99WO-JP06473.

XX 20-NOV-1998; 98JP-0347806.

XX (FUSO ) FUSO PHARM IND LTD.

XX Uemura H, Okui A, Kominami K, Yamaguchi N, Mitsui S;

XX WPI; 2000-400058/34.

XX N-PSDB; AAA61734.

XX Serine proteases BSSP5, useful in detecting homologs, mutants and

PT polymorphic variants as markers for diagnosis of e.g. Alzheimer's  
PT disease, epilepsy, cancer and inflammation, using blood, urine,  
XX pancreas or other tissues  
XX  
XX Claim 3; Page 55-56; 70pp; Japanese.  
XX  
CC The invention relates to novel serine proteases designated BSSP5  
CC (AAB11710-B11711), and to nucleic acids encoding them (AAA61733-A61734).  
CC The invention also relates to vectors and transformants comprising BSSP5  
CC nucleic acids; transgenic animals in which the expression level of BSSP5  
CC can be varied; and an mBSSP5 knockout mouse. The invention additionally  
CC encompasses anti-BSSP5 antibodies and methods of production of such  
CC antibodies, methods of BSSP5 detection using the antibodies, and the  
CC use of BSSP5 proteins or fragments as diagnostic markers for certain  
CC medical conditions, e.g., pancreatitis. A method for detecting  
CC pancreatitis comprising measuring BSSP5 concentration in the blood or  
CC urine, and a pancreatitis diagnostic agent containing an anti-BSSP5  
CC antibody is also disclosed. Nucleotides encoding BSSP5 were initially  
CC isolated in a human brain cDNA library using degenerate PCR primers  
CC (AAA61744-A61745) based on conserved regions of serine proteases. The  
CC BSSP5 serine proteases and nucleotides encoding them are useful in  
CC detecting homologues, mutants and polymorphic variants in biological  
CC samples (e.g., blood, urine, brain, prostate gland, placenta, testis,  
CC pancreas and spleen) as diagnostic markers for conditions such as  
CC Alzheimer's disease, epilepsy, cancer, inflammation, infertility,  
CC pancreatitis and prostatic hypertrophy. Sequence AAB11710 represents  
CC human BSSP5 (hBSSP5), and sequence AAB11711 represents murine BSSP5  
CC (mBSSP5).  
XX  
SQ Sequence 264 AA;  
  
Query Match 100.0%; Score 1418; DB 21; Length 264;  
Best Local Similarity 100.0%; Pred. No. 7.9e-112;  
Matches 264; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MLLSLTSLVLLGSGGCGVPAITPALSYNQRIVNGENAVPGSWPQVSLQDNTGFHC 60  
DB 1 MLLSLTSLVLLGSGGCGVPAITPALSYNQRIVNGENAVPGSWPQVSLQDNTGFHC 60  
  
QY 61 GGLISPNWVYTAACHQVTPGRHFVVLGEYDRSSNAEPQVLSIARAITHPNNTMNN 120  
DB 61 GGLISPNWVYTAACHQVTPGRHFVVLGEYDRSSNAEPQVLSIARAITHPNNTMNN 120  
  
QY 121 DLTLKLPARTYAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQVWLP 180  
DB 121 DLTLKLPARTYAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQVWLP 180  
  
QY 181 LVTNOCROYWGARIITDAMICAGSGGCGDGGPLVCQKNTWVLIGIVSWGTKNCN 240  
DB 181 LVTNOCROYWGARIITDAMICAGSGGCGDGGPLVCQKNTWVLIGIVSWGTKNCN 240  
  
QY 241 IQAPAMYTRVSKFSTWVNOVMAYN 264  
DB 241 IQAPAMYTRVSKFSTWVNOVMAYN 264  
  
RESULT 2  
AAB11710  
ID AAB11710 standard; Protein; 264 AA.  
XX AAB11710;  
AC AAB11710;  
XX  
DT 23-OCT-2000 (first entry)  
DE Human serine protease BSSP5 (hBSSP5) SEQ ID NO:2.  
XX BSSP5; serine protease; human; hBSSP5; mouse; mBSSP5; brain;  
KW diagnostic marker; antibody; transgenic animal; Alzheimer's disease;  
KW epilepsy; cancer; inflammation; infertility; pancreatitis;  
XX prostatic hypertrophy.  
OS Homo sapiens.  
XX

PN WO200031243-A1.  
XX  
XX 02-JUN-2000.  
XX  
PF 19-NOV-1999; 99WO-JP06473.  
XX  
XX 20-NOV-1998; 98JP-0347806.  
XX  
XX (FUSO ) FUSO PHARM IND LTD.  
XX  
XX Uemura H, Okui A, Kominami K, Yamaguchi N, Mitsui S;  
PI WPI; 2000-400058/34.  
XX  
XX N-PSDB; AAA61733.  
XX  
XX Serine proteases BSSP5, useful in detecting homologs, mutants and  
PT polymorphic variants as markers for diagnosis of e.g. Alzheimer's  
PT disease, epilepsy, cancer and inflammation, using blood, urine,  
XX pancreas or other tissues  
XX  
XX Claim 1; Page 51-52; 70pp; Japanese.  
XX  
CC The invention relates to novel serine proteases designated BSSP5  
CC (AAB11710-B11711), and to nucleic acids encoding them (AAA61733-A61734).  
CC The invention also relates to vectors and transformants comprising BSSP5  
CC nucleic acids; transgenic animals in which the expression level of BSSP5  
CC can be varied; and an mBSSP5 knockout mouse. The invention additionally  
CC encompasses anti-BSSP5 antibodies and methods of production of such  
CC antibodies, methods of BSSP5 detection using the antibodies, and the  
CC use of BSSP5 proteins or fragments as diagnostic markers for certain  
CC medical conditions, e.g., pancreatitis. A method for detecting  
CC pancreatitis comprising measuring BSSP5 concentration in the blood or  
CC urine, and a pancreatitis diagnostic agent containing an anti-BSSP5  
CC antibody is also disclosed. Nucleotides encoding BSSP5 were initially  
CC isolated in a human brain cDNA library using degenerate PCR primers  
CC (AAA61744-A61745) based on conserved regions of serine proteases. The  
CC BSSP5 serine proteases and nucleotides encoding them are useful in  
CC detecting homologues, mutants and polymorphic variants in biological  
CC samples (e.g., blood, urine, brain, prostate gland, placenta, testis,  
CC pancreas and spleen) as diagnostic markers for conditions such as  
CC Alzheimer's disease, epilepsy, cancer, inflammation, infertility,  
CC pancreatitis and prostatic hypertrophy. Sequence AAB11710 represents  
CC human BSSP5 (hBSSP5), and sequence AAB11711 represents murine BSSP5  
CC (mBSSP5).  
XX  
SQ Sequence 264 AA;  
  
Query Match 89.0%; Score 1262; DB 21; Length 264;  
Best Local Similarity 86.0%; Pred. No. 1.2e-98;  
Matches 227; Conservative 25; Mismatches 12; Indels 0; Gaps 0;  
  
QY 1 MLLSLTSLVLLGSGGCGVPAITPALSYNQRIVNGENAVPGSWPQVSLQDNTGFHC 60  
DB 1 MLLSLTSLVLLGSGGCGVPAITPALSYNQRIVNGENAVPGSWPQVSLQDNTGFHC 60  
  
QY 61 GGLISPNWVYTAACHQVTPGRHFVVLGEYDRSSNAEPQVLSIARAITHPNNTMNN 120  
DB 61 GGLISPNWVYTAACHQVTPGRHFVVLGEYDRSSNAEPQVLSIARAITHPNNTMNN 120  
  
QY 121 DLTLKLPARTYAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQVWLP 180  
DB 121 DLTLKLPARTYAQVSPVCLASTNEALPSGLTCVTTGWRISGVGNVTPARLQVWLP 180  
  
QY 181 LVTNOCROYWGARIITDAMICAGSGGCGDGGPLVCQKNTWVLIGIVSWGTKNCN 240  
DB 181 LVTNOCROYWGARIITDAMICAGSGGCGDGGPLVCQKNTWVLIGIVSWGTKNCN 240  
  
QY 241 IQAPAMYTRVSKFSTWVNOVMAYN 264  
DB 241 VRAPAVYTRVSKFSTWVNOVIAYN 264  
  
RESULT 3





CC over-expressed in ovarian tumours. TADG-15 protein or its fragments of  
 CC 9-20 residues that lack TADG-15 protease activity are useful for  
 CC vaccinating an individual against TADG-15, having, suspected of having or  
 CC at risk of getting cancer. Furthermore, the TADG-15 gene can be used as a  
 CC diagnostic or therapeutic target in cancer. The present sequence was used  
 CC in a sequence homology alignment with the catalytic domain of TADG-15.  
 XX  
 SQ Sequence 231 AA;

Query Match 53.3%; Score 755.5; DB 22; Length 231;  
 Best Local Similarity 58.2%; Pred. No. 5.8e-56;  
 Matches 135; Conservative 35; Mismatches 61; Indels 1; Gaps 1;  
 QY 33 RIVGENAVPGSWPQVSLQDNTGFHFCGGLISPNWVTAHQCQVTPGRHFVVLGEYDR 92  
 DB 1 RIVGENDAVPGSWPQVSLQDNTGFHFCGGLISEDWVTAHCGVRTS-DVVVAGEEDQ 59  
 QY 93 SSNAEPVOVLRIARITHPNKNANTMNDLTLKLPASRYTAQVSPVCLASTNEALPSG 152  
 DB 60 GSDEENIQVLKIAKVFKNPKFSILTAVNDITLLKLPAREFOTVSACVCLPSADDDPAG 119  
 QY 153 LTCVTTGGRISGVGNVTPARLQVPLVTVNOCROYWGARIITDAMICAGSGASCQ 212  
 DB 120 TLCATTGKTKYKANKTPDKLQQAALLSLNAECKKSWGRRIITDVMICAGSGVSCMG 179  
 QY 213 DSGGPLVCQKNTWVLIGIVSWGTKNCNIOAPAMTYTRVSKFSTWVNOVMAYN 264  
 DB 180 DSGGPLVCQKNTWVLIGIVSWGSDTCTSSPGYARVTKLIPVWQKILAN 231

RESULT 7  
 AAP61724  
 ID AAP61724 standard; Protein; 269 AA.  
 XX AAP61724;  
 AC AAP61724;  
 XX 23-OCT-1991 (first entry)  
 DT Porcine elastase II.  
 DE Colibacillus; yeast.  
 XX Sus scrofa.  
 OS JP61192289-A.  
 PN 26-AUG-1986.  
 PD 22-FEB-1985; 85JP-0034050.  
 PF 22-FEB-1985; 85JP-0034050.  
 PR (KIRI ) KIRIN BREWERY KK.  
 PA WPI; 1986-262895/40.  
 XX N-PSDB; AAN60919.  
 DR Biological prepn. of pig elastase II - using DNA with base  
 PT sequence coding for specified aminoacid sequence.  
 XX Claim 1; Fig 1; 11pp; Japanese.  
 PS The porcine elastase product may be efficiently expressed from a  
 CC transformed host such as colibacillus or yeast, yielding the product  
 CC more efficiently than spleen extraction.  
 CC See also J61192288.  
 XX  
 SQ Sequence 269 AA;  
 Query Match 37.98%; Score 537; DB 7; Length 269;  
 Best Local Similarity 43.1%; Pred. No. 2e-37;  
 Matches 118; Conservative 48; Mismatches 92; Indels 16; Gaps 11;

QY 2 LLSLTLTLVLLGSSWCGVPAITPALSYNORIVNGENAVPGSWPQVSLQ-DNTG--FH 58  
 DB 1 MIRALLSTLVAG-ALSCGLPANLPQL--PRVVGEDARPNSWPQVSLQYDSSGQWRH 56  
 QY 59 FCGSLSPNNWVTAHQCQVTPGRHFVVLGEYDRSSNAEPVOVLRIARITHPNKNANTM 118  
 DB 57 TCGTGLVDQSWLVAHCAHCISSTRYRVVLRHSLSTNEPGLAVKVKSLVLRHQMNSQL 116  
 QY 119 N--NDLTLKLPASRYTAQVSPVCLASTNEALPSGLTCVTTGGRISGVGNVTPARLQ 176  
 DB 117 SKGNDIALKLPASRYLTKLQGLCLPAAGSILPNYVYVTVGRLTNG-ASPDIIQQ 175  
 QY 177 VVLPLVTVNOCQ--YWGARIITDAMICAGSG-ASSCGDGGPLVCQKGN-TWVLIGIV 232  
 DB 176 GQLLVVDYATCSKPCWGSSTVKTNNICAGDGIISCCNGDGGPLNCOGANGQVHGIV 235  
 QY 233 SWGTK-NCN-TOAPAMTYTRVSKFSTWVNOVMAYN 264  
 DB 236 SFGSLGNCNYHKPSVETRVSNYIDWINSVIANN 269  
 RESULT 8  
 AAP70758  
 ID AAP70758 standard; Protein; 269 AA.  
 XX AAP70758;  
 AC AAP70758;  
 DT 19-APR-1991 (first entry)  
 XX Pig pancreas elastase-2.  
 DE Lipoprotein metabolism; lipid.  
 KW Sus scrofa.  
 OS Key Location/Qualifiers  
 FH Misc-difference 2..16  
 FT /note= "May be absent"  
 XX JP62000276-A.  
 PN 06-JAN-1987.  
 PD 25-JUN-1985; 85JP-0138494.  
 PF 25-JUN-1985; 85JP-0138494.  
 PR (SANY ) SANKYO KK.  
 XX WPI; 1987-040875/06.  
 DR N-PSDB; AAN71121.  
 XX Pancreas elastase prodn. used to improve lipoprotein metabolism -  
 PT comprises isolating RNA coding elastase, synthesising single and  
 PT double chain cDNA and introducing recombinant into host  
 XX Disclosure; Page 480; 18pp; Japanese.  
 PS Sequence encoding pancreas elastase may be used to transform an  
 CC expression system. The product is useful in rejuvenating the  
 CC elasticity of the arterial wall, and improving unusual serum lipid  
 CC levels and lipoprotein metabolism.  
 XX  
 SQ Sequence 269 AA;  
 Query Match 37.98%; Score 537; DB 8; Length 269;  
 Best Local Similarity 43.1%; Pred. No. 2e-37;  
 Matches 118; Conservative 48; Mismatches 92; Indels 16; Gaps 11;

QY 2 LLSLTLTLVLLGSSWCGVPAITPALSYNORIVNGENAVPGSWPQVSLQ-DNTG--FH 58  
 DB 1 MIRALLSTLVAG-ALSCGLPANLPQL--PRVVGEDARPNSWPQVSLQYDSSGQWRH 56

```

QY 59 FCGGSLISPNWVYVTAHCOVTPGRHFVVLGEYDRSSNAEPVOVLSTARAITHPNNNANTM 118
DB 57 TCGGTLVDQSWYLTAAHCITSSRTYRVVLGRHSLSTNEPGLAVKSVLKVHQQDWNNSOL 116
QY 119 N--NDLTLLKASPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGCVGNVTPARLQO 176
DB 117 SGNNDIALKLKASPVSLSLTDKIQLGCLPAAAGTILPNNYCVYVWGRLQTNNG--ASPDILQO 175
QY 177 VVPLVTVNOCRQ--YWGARIIDAMICAGGSG--ASSCGDGGGGLVPCQKGN--TWVLIGIV 232
DB 176 GQLLVVDYATCSKPGWGSTVTKTMICAGGDCGISSCNGDGGGLNCGQANGQVQVHGIV 235
QY 233 SWGTK--NCN--IQAPAMYTRVSKFTWVNOVMAYN 264
DB 236. SFGSSILGCNYHKPSVFTRVSNYIDWINSVIANN 269

RESULT 9
AAR29621
ID AAR29621 standard; Protein; 253 AA.
XX
AC AAR29621;
XX
DT 30-APR-1993 (first entry)
XX
DE Porcine pancreatic elastase III.
XX
KW Pig; swine; bile acid; gall bladder; bile acid secretion promoters;
KW liver function improvers; ss.
XX
OS Sus scrofa domestica.
XX
PN JP04325090-A.
XX
PD 13-NOV-1992.
XX
PF 23-APR-1991; 91JP-0092069.
XX
PR 23-APR-1991; 91JP-0092069.
XX
PA (SANY ) SANKYO CO LTD.
XX
DR WPI; 1992-428826/52.
DR N-PSDB; AAQ31724.
XX
PT Pig pancreas elastase protein - used in bile acid secretion
PT promoters and liver function improvers
XX
PS Claim 1; Page 8; 10pp; Japanese.
XX
CC This sequence represents porcine pancreatic elastase III. It was
CC expressed in E. coli YA21 using the expression vector pELE001.
CC It may be used in bile acid secretion promoters and liver function
CC improvers.
XX
SQ Sequence 253 AA;

Query Match 36.9%; Score 523.5; DB 13; Length 253;
Best Local Similarity 43.0%; Pred. No. 2.5e-36;
Matches 110; Conservative 43; Mismatches 86; Indels 17; Gaps 10;

QY 19 CGVPALTTPALSTNQRIVNGENAVPGSWPMQVSLQ--DNTGF--HFCCGSLISPNWVYVTAH 75
DB 1 CGRPSYNPA----SRVNGEDAVPYSPWPMQVSLQYEKNGVFOHTCGGSLIAPDNLVITAGH 56
QY 76 CQVTPGRHFVVLGEYDRSSNAEPVQLSI--ARAITHPNNNANTM--NDLTLLKASPARY 131
DB 57 CISSSLTYQVVLGEYDRSNEGFEQVPIPNAGDLFVHPRWNSNCVSCGNDIALVKLSRSA 116
QY 132 RYTAQVSPVCLASTNEALPSGLTCVTTGWRISGCVGNVTPARLQOVLPLVTVNOCRQ-- 189
DB 117 QLGDKVQLACLPPAGDILPNDTPFCYISGWGRLYTNGPL--PDKLOQALLPLPVDYQHCWKD 175

```

```

QY 190 YGARITDAMICAGGSGASSCGDGGGLVPCQKGN--TWVLIGIVSW--GFKNCN--IQAPAM 246
DB 176 WNGSTVKOTMVCAGDIRSGCNGSGGGLNCPAADGSMQVHGVTSEVSAYGCTNLTKKPTV 235
QY 247 YTRVSKFTWVNOVMA 262
DB 236 FVRTSAFDWIEEIIA 251

RESULT 10
AAR60062
ID AAR60062 standard; Protein; 269 AA.
XX
AC AAR60062;
XX
DT 06-SEP-1991 (first entry)
XX
DE Sequence of human pancreatic elastase IIA encoded on pH2E2.
XX
KW Enzyme; serum lipoprotein metabolism.
XX
OS Homo sapiens.
XX
PN EP198645-A.
XX
PD 22-OCT-1986.
XX
PF 07-APR-1986; 86EP-0302557.
XX
PR 23-OCT-1985; 85JP-0236686.
PR 05-APR-1985; 85JP-0072308.
PR 27-APR-1985; 85JP-0091986.
PR 26-JUL-1985; 85JP-0163964.
PR 02-DEC-1985; 85JP-0271128.
XX
PA (SANY ) SANKYO KK.
XX
PI Takiguchi Y, Tani T, Kawashima I, Eurukawa H, Ohmine T;
PI Ohsumi J;
XX
DR WPI; 1986-280300/43.
DR N-PSDB; AAN60076.
XX
PT New genetically engineered human pancreatic elastase - obt'd.
PT using hosts modified DNA coding for enzyme
XX
PS Example; Page 14-15; 45pp; English.
XX
CC The genetically engineered prod. can eliminate the dependency on
CC human pancreas supplies for the elastase, and avoids antibody
CC formation and possibility of anaphylaxis using porcine elastase.
XX
SQ Sequence 269 AA;

Query Match 36.3%; Score 515; DB 7; Length 269;
Best Local Similarity 41.6%; Pred. No. 1.4e-35;
Matches 114; Conservative 47; Mismatches 97; Indels 16; Gaps 10;

QY 2 LLSLTLSVLSSGCGVPAITPALSINQRIVNGENAVPGSWPMQVSLQDNTG---FH 58
DB 1 MIRTLLSTLVAG--ALSCGDPTYP---YVTRVVGGEARPNPWPQVSLQYSSNGKWT 56
QY 59 FCGGSLISPNWVYVTAHCOVTPGRHFVVLGEYDRSSNAEPVOVLSTARAITHPNNNANTM 118
DB 57 TCGGSLIANSWLTAAHCITSSRTYRVVLGRHNLVVAESGSLAVSVSKIVVHKDWNNOI 116
QY 119 N--NDLTLLKASPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGCVGNVTPARLQO 176
DB 117 SGNNDIALKLKASPVSLSLTDKIQLACLPAPAGTILPNNYCVYVWGRLQTNNGAV--PDVLQO 175
QY 177 VVPLVTVNOCRQ--YWGARIIDAMICAGGSGA--SSCGDGGGLVPCQKGN--TWVLIGIV 232
DB 176 GRLLVVDYATCSSAAWGGSSVKTSMICAGDGVVSSCNGDGGGLNCPQASDGRWQVHGIV 235

```





```

Db 117 SKGNDIALLKLANPVSITDKIQIACLPAGTILPNPYPCYVTGWRLOTNGAV-PDVLOQ 175
QY 177 WVLPLVTNOCRO--YNGARITDAMICAGSGA-SSCGSDSGGLVCQKQNTWVLI-GIV 232
Db 176 GRLLVVDYATCSSAWGSSVKTSMICAGDGVISSCNGSDGGPLNCQASDGRCEVHGIV 235
QY 233 SWGTK-NCN-IQAPAMYTRVSKFSTWINQWYN 264
Db 236 SFGSLGCGNYHKPVSFTRVSNIDWINSIANN 269

RESULT 13
AAR90683
ID AAR90683 standard; Protein: 268 AA.
XX
XX AAR90683;
XX
DT 11-NOV-1996 (first entry)
XX
DE Human caldecrin contg. preprosequence.
XX
KW caldecrin; rat; human; calcium serum lowering agent;
KW recombinant production; preprosequence.
XX
OS Rattus sp.
XX
FH Key
FT Peptide 1..13
FT /label= pre-pro-sequence
FT Peptide 16..29
FT /label= pro-sequence
FT Protein 30..268
FT /label= mature_protein
FT Misc-difference 80
FT /label= Trp, Arg
XX
PN W09600287-A.
XX
XX 04-JAN-1996.
XX
XX 23-JUN-1995; 95W0-JP01268.
XX
XX 08-MAR-1995; 95JP-0074676.
XX 24-JUN-1994; 94JP-0164898.
XX
XX (CHUS ) CHUGAI SEIYAKU KK.
XX
XX Akiyama M, Noikura T, Saheki T, Tomomura A;
XX
XX WPI; 1996-068871/07.
XX N-PSDB; AAT15536.
XX
XX DNA encoding rat and human caldecrin and related vectors - for prodn.
XX of caldecrin for use as serum calcium lowering agent
XX
XX Claim 11; Page 26-27; 43pp; Japanese.
XX
XX Human caldecrin preprosequence is encoded by AAT15536. Expression
XX vectors contg. the DNA can be used for the recombinant production of
XX caldecrin from transformant host cells. The caldecrin is used as a
XX serum calcium lowering agent.
XX
SQ Sequence 268 AA;
Query Match 35.0%; Score 496.5; DB 17; Length 268;
Best Local Similarity 40.1%; Pred. No. 5.2e-34;
Matches 108; Conservative 46; Mismatches 102; Indels 13; Gaps 9;

QY 3 LLSLTLSLVLGSSGCGVPATPALSYNQRIYNGENAVPGSWPQVSLQ---DNTGFHF 59
Db 1 MLGITVLAALLACASCGVPSPFPNLS--ARVVGEDARHSPWPQISLQYLKNDTWRT 58
QY 60 CGSGLISPNWVVTAAHCQVTPGPHFVVLGEYD-RSSNAEPVQVLSIARAITHPNWNANTM 118

```

```

Db 59 CGGTLASNPVLTAACISNTXYRVAVGKNNLEVEDEGSLFVGVDTHVHRKNALLL 118
QY 119 NNDLTLLKLASPARYTAQVSPVCLASTNEALPSGLTCVTTGWRISGNGVNTPARLQOVV 178
Db 119 RNDIALIKLAHVLSLSDTIOVACLPEKDSLLPKDYPCYVTGWRLTNGPIAD-KLOOGL 177
QY 179 LPLVTYNOCRQ--YMGARITDAMICAGSGA-SSCGSDSGGLVCQKGN-TWVLIGIVSW 234
Db 178 QPVVDHATCSRIDWGFVRVKKTWVCAGDGVISACNGDSGGPLNCOLENGSWEVFGIVSF 237
QY 235 GT-KNCNI-QAPAMYTRVSKFSTWINQVM 261
Db 238 GSRRGCGNTRKKPVYTRVSAYIDWINEKM 266

RESULT 14
AAP60058
ID AAP60058 standard; Protein: 252 AA.
XX
XX AAP60058;
XX
DT 06-SEP-1991 (first entry)
XX
DE Sequence of human pancreatic elastase IIA.
XX
KW Enzyme; serum lipoprotein metabolism.
XX
OS Homo sapiens.
XX
PN EP198645-A.
XX
PD 22-OCT-1986.
XX
PF 07-APR-1986; 86EP-0302557.
XX
XX 23-OCT-1985; 85JP-0236686.
XX 05-APR-1985; 85JP-0072308.
XX 27-APR-1985; 85JP-0091986.
XX 26-JUL-1985; 85JP-0163964.
XX 02-DEC-1985; 85JP-0271128.
XX
XX (SANY ) SANKYO KK.
XX
XX Takiguchi Y, Tani T, Kawashima I, Eurukawa H, Ohmine T;
XX Ohsumi J;
XX
XX WPI; 1986-280300/43.
XX N-PSDB; AAN60072.
XX
XX New genetically engineered human pancreatic elastase - obt'd.
XX using hosts modified DNA coding for enzyme
XX
XX Claim 2; Page 3-4; 45pp; English.
XX
XX The genetically engineered prod. can eliminate the dependency on
XX human pancreas supplies for the elastase, and avoids antibody
XX formation and possibility of anaphylaxis using porcine elastase.
XX
SQ Sequence 252 AA;
Query Match 34.7%; Score 492; DB 7; Length 252;
Best Local Similarity 42.4%; Pred. No. 1.2e-33;
Matches 109; Conservative 41; Mismatches 91; Indels 16; Gaps 10;

QY 19 CGVPATPALSYNQRIYNGENAVPGSWPQVSLQDNTG---PHFCGSLISPNWVVTAAH 75
Db 1 CGDPTYP---YVTRVVGGEARPNSPWQVSLQYSSNGKWYHTCGSGLIANSWVLTAH 57
QY 76 CQVTPGRHFVVLGEYDRSSNAEPVQVLSIARAITHPNWNANTM--NDLTLLKLASPARY 133
Db 58 C1SSSRTRYGLGRHNLVVAESGSLAVSYSKIIVVHKDWSNQISKGNIDIALKLANPVSL 117

```

Db	677	VITDKVIPACLPSPNVVYVADRTTECFITGWGETGTGFA--GLLKEAQLPVNIENKVCNRVE	734
QY	191	-WGARIITDAMICAG--CGSASSCQDGGGGLVQCQKNTWVLIGIVSWGKNCNIQAPMY	247
Db	735	FLNVRVOSTELCAGHLAGGTDSCQDGGGLVCFEKDKYILOQVTSWGLGCARPKNPGVY	794
QY	248	TRVSKFSTWINOVNAYN	264
Db	795	VRVSRFVTWIEGVNRRN	811

Search completed: December 20, 2002, 15:16:41  
Job time : 42.6957 secs

[illegible]



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OK protein - protein search, using sw model

Run on: December 20, 2002, 15:16:49 ; Search time 7.30435 Seconds  
(without alignments)  
530.147 Million cell updates/sec

Title: us-09-856-319b-2\_COPY\_1\_231

Perfect score: 1221

Sequence: 1 MLLSLTLVLGSSWGGC.....GDSGGPLVCQKGTWVLIGI 231

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 106657 seqs, 16763532 residues

Total number of hits satisfying chosen parameters: 106657

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	996	81.6	192	10	US-09-925-297-529
2	684.5	56.1	263	10	US-09-888-615-96
3	427.5	35.0	269	10	US-09-925-297-576
4	420.5	34.4	270	10	US-09-923-779-152
5	420.5	34.4	273	10	US-09-925-297-695
6	410.5	33.6	343	10	US-09-948-094-2
7	408.5	33.5	453	9	US-09-978-295A-69
8	408.5	33.5	453	9	US-09-978-697-69
9	408.5	33.5	453	9	US-09-978-192A-69
10	408.5	33.5	453	12	US-10-052-586-64
11	408	33.4	248	10	US-09-925-301-1017
12	408	33.4	327	10	US-09-804-156-16
13	408	33.4	327	10	US-09-946-633-8
14	408	33.4	327	10	US-09-888-615-103
15	397	32.5	791	9	US-09-967-386-1
16	397	32.5	810	10	US-09-946-893-2
17	389.5	31.9	812	9	US-09-335-325-1
18	389.5	31.9	812	10	US-09-788-142-1
19	389.5	31.9	812	10	US-09-761-120-1

20	389.5	31.9	812	10	US-09-873-676-81	Sequence 81, Appl
21	388.5	31.8	146	10	US-09-925-297-643	Sequence 643, App
22	382	31.3	235	10	US-09-988-975A-6	Sequence 6, Appl1
23	382	31.3	279	9	US-09-858-332-15	Sequence 15, Appl
24	379	31.0	393	9	US-10-012-896-934	Sequence 934, App
25	379	31.0	393	10	US-09-759-143-934	Sequence 934, App
26	379	31.0	393	10	US-09-780-669-934	Sequence 934, App
27	379	31.0	393	10	US-09-822-827-934	Sequence 934, App
28	379	31.0	492	9	US-10-012-896-932	Sequence 932, App
29	379	31.0	492	10	US-09-759-143-932	Sequence 932, App
30	379	31.0	492	10	US-09-780-669-932	Sequence 932, App
31	379	31.0	492	10	US-09-822-827-932	Sequence 932, App
32	378.5	31.0	638	9	US-09-808-602-102	Sequence 102, App
33	378	31.0	249	9	US-09-961-721-5	Sequence 5, Appl1
34	378	31.0	251	9	US-09-961-721-4	Sequence 4, Appl1
35	377.5	30.9	370	10	US-09-888-615-101	Sequence 101, App
36	376	30.8	384	9	US-09-981-353-23	Sequence 23, Appl
37	375.5	30.8	226	9	US-09-961-721-6	Sequence 6, Appl1
38	375.5	30.8	226	9	US-10-045-367A-5	Sequence 5, Appl1
39	375.5	30.8	226	10	US-09-910-151-6	Sequence 6, Appl1
40	374	30.6	492	9	US-10-012-896-895	Sequence 895, App
41	374	30.6	492	10	US-09-759-143-895	Sequence 895, App
42	374	30.6	492	10	US-09-780-669-895	Sequence 895, App
43	374	30.6	492	10	US-09-879-792-14	Sequence 14, Appl
44	374	30.6	492	10	US-09-822-827-895	Sequence 895, App
45	370	30.3	223	10	US-09-910-071-14	Sequence 14, Appl

## ALIGNMENTS

RESULT 1  
US-09-925-297-529  
; Sequence 529, Application US/09925297  
; Patent No. US20020081659A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA105  
; CURRENT APPLICATION NUMBER: US/09/925,297  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05989  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 928  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 529  
; LENGTH: 192  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-925-297-529

Query Match	81.6%	Score 996;	DB 10;	Length 192;
Best Local Similarity	99.5%;	Pred. No. 8.4e-83;		
Matches 189;	Conservative	0;	Mismatches 1;	Indels 0; Gaps 0;
QY	5	SUTLSLVLLGSSWGGCIPALPAKLSFSQRIYNGENAVLGSWFWQVSLQDSSGFHFCGSSL	64	
Db	1	SUTLSLVLLGSSWGGCIPALPAKLSFSQRIYNGENAVLGSWFWQVSLQDSSGFHFCGSSL	60	
QY	65	ISQSWVTAAHCNVSFGRHFVVLGVEDRSSNAEPQLQVLSRAITHPSWNSTTMMNDVLT	124	
Db	61	ISQSWVTAAHCNVSFGRHFVVLGVEDRSSNAEPQLQVLSRAITHPSWNSTTMMNDVLT	120	
QY	125	LKLSPAQYTTTRISPVCLASSNEALTEGLTCVTTGGRSLSGVGNVTPAHLQVALPLTV	184	
Db	121	LKLSPAQYTTTRISPVCLASSNEALTEGLTCVTTGGRSLSGVGNVTPAHLQVALPLTV	180	
QY	185	NOCROYWSSS	194	
Db	181	NOCROYWSSS	190	



Qy 231 I 231

Db 235 V 235

#### RESULT 5

US-09-925-297-695  
; Sequence 695, Application US/09925297  
; Patent No. US20020081659A1

#### GENERAL INFORMATION:

; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA105

; CURRENT APPLICATION NUMBER: US/09/925,297

; CURRENT FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/05989

; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270

; PRIOR FILING DATE: 1999-03-12

; NUMBER OF SEQ ID NOS: 928

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 695

; LENGTH: 273

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (27)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE

; LOCATION: (28)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE

; LOCATION: (34)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-925-297-695

Query Match 34.4%; Score 420.5; DB 10; Length 273;  
Best Local Similarity 40.7%; Pred. No. 9.2e-31;  
Matches 98; Conservative 34; Mismatches 92; Indels 17; Gaps 8;

Qy 1 MLLSLTSLVLLGSSWCGGIPAKIPALSFQRIYNGENAVLGSWPMQVLSQ-DSSG--F 57

Db 5 MLRLSSLLLVAVASGYG-----PPSSXXSRVVGEDAVPYSPWQVLSQYKSGSFY 58

Qy 58 HFCGSLISQSWVYTAHCNVSPGRHFVVLGEYDRSSNAEPLQVLSVS--RAITHPSWNS 115

Db 59 HFCGSLIAPQWVYTAHCNVSPGRHFVVLGEYDRSSNAEPLQVLSVS--RAITHPSWNS 118

Qy 116 TTM--NNDVTLKLKSPAQYTRISPVCLASSNEALTEGLTCVTTGWGRLSGVGNVTPAH 173

Db 119 SCVACGNDIALIKLSRAQLGDAVOLASLPAGDILPNKTPCYITGWGRLYTNGLP-PDK 177

Qy 174 LQOVALPLVTNQCRQ--YWDSITDSMICAGGASSCGSGGPLVC-QKGNWTWVLIG 230

Db 178 LQOARLPVVDYKHCSSWNNWGSVTKKTVACAGGYIRSCNGDSGGPLNCPTEDGGWQVHG 237

Qy 231 I 231

Db 238 V 238

#### RESULT 6

US-09-948-094-2

; Sequence 2, Application US/09948094

; Patent No. US20020090625A1

#### GENERAL INFORMATION:

; APPLICANT: The Brigham and Women's Hospital, Inc.

; APPLICANT: Mok, Samuel

; APPLICANT: Wong, Kwong-kwok

; TITLE OF INVENTION: Methods of Detecting Cancer Based on Prostatein

; FILE REFERENCE: 81994/282423

; CURRENT APPLICATION NUMBER: US/09/948,094

; CURRENT FILING DATE: 2001-09-07

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 2

; LENGTH: 343

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-948-094-2

Query Match 33.6%; Score 410.5; DB 10; Length 343;

Best Local Similarity 36.3%; Pred. No. 9.8e-30;

Matches 91; Conservative 38; Mismatches 89; Indels 33; Gaps 8;

Qy 4 LSLTSLVLLGSSWG-----CGIPAKIPALSFQRIYNGENAVLGSWPMQVLSQDSSG 56

Db 15 VAILLYLGLLRSGTGAEGAEAPCGVAP-----QARITGSSAVAGQMPQVSI-TYEG 66

Qy 57 FHFCGSLISQSWVYTAHCNVSPGRHF-----VVLGEYDRSSNAEPLQVLSVSRATHP 111

Db 67 VHVCGSLVSEQWVLSAAHC--FPSEHHKEAYEVKLGHAQLDSYSEDAKVSTLKDIIPHP 124

Qy 112 SWNSTMMNDVTLKLKSPAQYTRISPVCLASSNEALTEGLTCVTTGWGRLS-GVGNVT 170

Db 125 SYLQEGSGDIALQLSRPITESRVIRPCLPAANASFPNGLHCTVTGNGHVAPSLSLT 184

Qy 171 PAHLOQVALPLVTNQCRQYWDSS-----ITDSMICAG--GAGASSCGDSGGPLVC 220

Db 185 PKPLOOLEVPLISRETCLYNIDAKPEHPFVQEDMVCAGYVEGKGKACOGDSGGPLSC 244

Qy 221 QKGNWTWVLIGI 231

Db 245 PVEGLWYLTGI 255

#### RESULT 7

US-09-978-295A-69

; Sequence 69, Application US/09978295A

; Patent No. US20020156006A1

#### GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Baker Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan

; APPLICANT: Ferrara, Napoleon

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Kuo, Sophia S.

; APPLICANT: Napier, Mary A.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Shelton, David L.

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; TITLE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: P2630PIC11

; CURRENT APPLICATION NUMBER: US/09/978,295A

; CURRENT FILING DATE: 2001-10-15

; PRIOR APPLICATION NUMBER: 09/918585

1 PRIOR FILING DATE: 2001-07-30  
2 PRIOR APPLICATION NUMBER: 60/062250  
3 PRIOR FILING DATE: 1997-10-17  
4 PRIOR APPLICATION NUMBER: 60/064249  
5 PRIOR FILING DATE: 1997-11-03  
6 PRIOR APPLICATION NUMBER: 60/065311  
7 PRIOR FILING DATE: 1997-11-13  
8 PRIOR APPLICATION NUMBER: 60/066364  
9 PRIOR FILING DATE: 1997-11-21  
10 PRIOR APPLICATION NUMBER: 60/077450  
11 PRIOR FILING DATE: 1998-03-10  
12 PRIOR APPLICATION NUMBER: 60/077632  
13 PRIOR FILING DATE: 1998-03-11  
14 PRIOR APPLICATION NUMBER: 60/077641  
15 PRIOR FILING DATE: 1998-03-11  
16 PRIOR APPLICATION NUMBER: 60/077649  
17 PRIOR FILING DATE: 1998-03-11  
18 PRIOR APPLICATION NUMBER: 60/077791  
19 PRIOR FILING DATE: 1998-03-12  
20 PRIOR APPLICATION NUMBER: 60/078004  
21 PRIOR FILING DATE: 1998-03-13  
22 PRIOR APPLICATION NUMBER: 60/078886  
23 PRIOR FILING DATE: 1998-03-20  
24 PRIOR APPLICATION NUMBER: 60/078936  
25 PRIOR FILING DATE: 1998-03-20  
26 PRIOR APPLICATION NUMBER: 60/078910  
27 PRIOR FILING DATE: 1998-03-20  
28 PRIOR APPLICATION NUMBER: 60/078939  
29 PRIOR FILING DATE: 1998-03-20  
30 PRIOR APPLICATION NUMBER: 60/079294  
31 PRIOR FILING DATE: 1998-03-25  
32 PRIOR APPLICATION NUMBER: 60/079656  
33 PRIOR FILING DATE: 1998-03-26  
34 PRIOR APPLICATION NUMBER: 60/079664  
35 PRIOR FILING DATE: 1998-03-27  
36 PRIOR APPLICATION NUMBER: 60/079689  
37 PRIOR FILING DATE: 1998-03-27  
38 PRIOR APPLICATION NUMBER: 60/079663  
39 PRIOR FILING DATE: 1998-03-27  
40 PRIOR APPLICATION NUMBER: 60/079728  
41 PRIOR FILING DATE: 1998-03-27  
42 PRIOR APPLICATION NUMBER: 60/079786  
43 PRIOR FILING DATE: 1998-03-27  
44 PRIOR APPLICATION NUMBER: 60/079920  
45 PRIOR FILING DATE: 1998-03-30  
46 PRIOR APPLICATION NUMBER: 60/079923  
47 PRIOR FILING DATE: 1998-03-30  
48 PRIOR APPLICATION NUMBER: 60/080105  
49 PRIOR FILING DATE: 1998-03-31  
50 PRIOR APPLICATION NUMBER: 60/080107  
51 PRIOR FILING DATE: 1998-03-31  
52 PRIOR APPLICATION NUMBER: 60/080165  
53 PRIOR FILING DATE: 1998-03-31  
54 PRIOR APPLICATION NUMBER: 60/080194  
55 PRIOR FILING DATE: 1998-03-31  
56 PRIOR APPLICATION NUMBER: 60/080327  
57 PRIOR FILING DATE: 1998-04-01  
58 PRIOR APPLICATION NUMBER: 60/080328  
59 PRIOR FILING DATE: 1998-04-01  
60 PRIOR APPLICATION NUMBER: 60/080333  
61 PRIOR FILING DATE: 1998-04-01  
62 PRIOR APPLICATION NUMBER: 60/080334  
63 PRIOR FILING DATE: 1998-04-01  
64 PRIOR APPLICATION NUMBER: 60/081070  
65 PRIOR FILING DATE: 1998-04-08  
66 PRIOR APPLICATION NUMBER: 60/081049  
67 PRIOR FILING DATE: 1998-04-08  
68 PRIOR APPLICATION NUMBER: 60/081071  
69 PRIOR FILING DATE: 1998-04-08  
70 PRIOR APPLICATION NUMBER: 60/081195  
71 PRIOR FILING DATE: 1998-04-08  
72 PRIOR APPLICATION NUMBER: 60/081203  
73 PRIOR FILING DATE: 1998-04-09  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000



```

RESULT 8
US-09-978-697-69
US-978-69, Application US/09978697
Patent No. US20020169284A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnovers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey E.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tunas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transferred

```

/	PRIOR APPLICATION NUMBER: 60/084643
/	PRIOR FILING DATE: 1998-05-07
/	PRIOR APPLICATION NUMBER: 60/085339
/	PRIOR FILING DATE: 1998-05-13
/	PRIOR APPLICATION NUMBER: 60/085338
/	PRIOR FILING DATE: 1998-05-13
/	PRIOR APPLICATION NUMBER: 60/085323
/	PRIOR FILING DATE: 1998-05-13
/	PRIOR APPLICATION NUMBER: 60/085582
/	PRIOR FILING DATE: 1998-05-15
/	PRIOR APPLICATION NUMBER: 60/085700
/	PRIOR FILING DATE: 1998-05-15
/	PRIOR APPLICATION NUMBER: 60/085689
/	PRIOR FILING DATE: 1998-05-15
/	PRIOR APPLICATION NUMBER: 60/085579
/	PRIOR FILING DATE: 1998-05-15
/	PRIOR APPLICATION NUMBER: 60/085580
/	PRIOR FILING DATE: 1998-05-15
/	PRIOR APPLICATION NUMBER: 60/085573
/	PRIOR FILING DATE: 1998-05-15
/	PRIOR APPLICATION NUMBER: 60/085704
/	PRIOR FILING DATE: 1998-05-15
/	PRIOR APPLICATION NUMBER: 60/085697

Query Match  
Best Local Similarity  
Matches

33.5%;  
39.9%;  
83;

Score 408.5;  
Pred. No. 2.1e-29;  
Conservative 31;

DB 9;  
Indels 11;  
Gaps 3;

Length 453;

QY	30	FSORIYGENAVLGSPPWQVSLDDSGFHFCGGSLISQSXWVTAAHCNVYS---	PGRHFVV 86
.Db	213	YSSRIYGVGNMSSLSSQPWQASLO-FOGYHLCGGSVITPLWIITAAHCYVDLYLPKSWTIQ 271	:
QY	87	LGEYDRSSNAEPLQVLVSRAITHPSWNSTTMNDVTLTKLASPAOYVTRISPVCLASSN 146	:
.Db	272	VGLVSLLDNPAPSHL--VEKIVTHSYKPKRUGNDIALMKLAGLPTFNEMIQVCLPNSE 329	:
QY	147	EALTEGLTCVTTGWRLSGVGNYTPAHLOQVALPLVTVNQC--ROYWDSSTIDSMICAG-	203
.Db	330	ENFPDGKVCTSCWGATEDGGDASP-V-LNHAAVPLISNKICNRDYVGGIISPMLCAGY 388	:
QY	204	-GAGASCQGDSCGGPLVCQKGNFTWLIG 230	:
.Db	389	LTGGVDSCQGDSCGGPLVCQBERRLWKLVG 416	:

RESULT 9

US-09-978-192A-69

Sequence 69, Application US/09978192A

Patent No. US2002017753A1

GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi

APPLICANT: Baker Kevin P.

APPLICANT: Botstein, David

APPLICANT: Desnoyers, Luc

APPLICANT: Eaton, Dan

APPLICANT: Ferrara, Napoleon

APPLICANT: Filvaroff, Ellen

APPLICANT: Fong, Sherman

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerber, Hanspeter

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Grimaldi, J. Christopher

APPLICANT: Gurney, Austin L.

APPLICANT: Hillan, Kenneth J.

APPLICANT: Kljavin, Ivar J.

APPLICANT: Kuo, Sophia S.

APPLICANT: Napier, Mary A.

APPLICANT: Pan, James;

APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann

APPLICANT: Shelton, David L.

APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2630PIC9  
CURRENT APPLICATION NUMBER: US/09/978.192A  
CURRENT FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 05/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083742  
PRIOR FILING DATE: 1998-04-30  
PRIOR APPLICATION NUMBER: 60/084366  
PRIOR FILING DATE: 1998-05-05  
PRIOR APPLICATION NUMBER: 60/084414  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084441  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084639  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084640  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084598

; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084600  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084627  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084643  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/085339  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085338  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085323  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085582  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085700  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085689  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085579  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085580  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085573  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085704  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085697

Query Match 33.5%; Score 408.5; DB 9; Length 453;

Best Local Similarity 39.9%; Pred. No. 2.1e-29;  
Matches 83; Conservative 31; Mismatches 83; Indels 11; Gaps 6;

QY 30 FSRIVNGENAVLGSWPQVSLQDSGPHFCGSLISQSWVTAHCNVS---PGRHFV 86  
 DB 213 YSSRIVGGNMSLLSQWPQASLQ-FQYHLCGGSVITPLWITAAHCNVYDLYLPKSWTIQ 271  
 QY 87 LGEYDRSSNAELQVLSVSRATHPSWSTTMNDVTLKLKASPAQYTRISPVCLASN 146  
 DB 272 VGLVSLDNPASHL--VEKIVYHSKYKPKRLGNDIALMKLAGPLTFNEMIQVCLPNSE 329  
 QY 147 EALTEGLCTVTGWRGLSGVGNVTPAHLQOVALPLTVVNO--RQWDSSTDSMTACG- 203  
 DB 330 ENFPDGVKVCWTSGWATEDGGDASPY-LNHAAPVLSNKNICHRDVGYYGIISPSMLCAGY 388  
 QY 204 -GAGASSCGDGGPLVCKGNTWVLIG 230  
 DB 389 LTGGVDSGCGDGGPLVCQERRLKLVG 416

RESULT 10

US-10-052-586-64  
 ; Sequence 64, Application US/10052586  
 ; Patent No. US20020127584A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Chen, Jian  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Pan, James  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Watanabe, Colin K.  
 ; APPLICANT: Wood, William I.  
 ; APPLICANT: Zhang, Zemin  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE REFERENCE: P3430R1C1  
 ; CURRENT APPLICATION NUMBER: US/10/052, 586  
 ; CURRENT FILING DATE: 2002-01-15  
 ; PRIOR APPLICATION NUMBER: 60/059263  
 ; PRIOR FILING DATE: 1997-09-18

; PRIOR APPLICATION NUMBER: 60/059266  
 ; PRIOR FILING DATE: 1997-09-18  
 ; PRIOR APPLICATION NUMBER: 60/062250  
 ; PRIOR FILING DATE: 1997-10-17  
 ; PRIOR APPLICATION NUMBER: 60/063120  
 ; PRIOR FILING DATE: 1997-10-24  
 ; PRIOR APPLICATION NUMBER: 60/063121  
 ; PRIOR FILING DATE: 1997-10-24  
 ; PRIOR APPLICATION NUMBER: 60/063486  
 ; PRIOR FILING DATE: 1997-10-21  
 ; PRIOR APPLICATION NUMBER: 60/063540  
 ; PRIOR FILING DATE: 1997-10-28  
 ; PRIOR APPLICATION NUMBER: 60/063541  
 ; PRIOR FILING DATE: 1997-10-28  
 ; PRIOR APPLICATION NUMBER: 60/063544  
 ; PRIOR FILING DATE: 1997-10-28  
 ; PRIOR APPLICATION NUMBER: 60/063564  
 ; PRIOR FILING DATE: 1997-10-28  
 ; PRIOR APPLICATION NUMBER: 60/063734  
 ; PRIOR FILING DATE: 1997-10-29  
 ; PRIOR APPLICATION NUMBER: 60/063870  
 ; PRIOR FILING DATE: 1997-10-31  
 ; PRIOR APPLICATION NUMBER: 60/064103  
 ; PRIOR FILING DATE: 1997-10-31  
 ; PRIOR APPLICATION NUMBER: 60/065311  
 ; PRIOR FILING DATE: 1997-11-13  
 ; PRIOR APPLICATION NUMBER: 60/066120  
 ; PRIOR FILING DATE: 1997-11-21  
 ; PRIOR APPLICATION NUMBER: 60/066466  
 ; PRIOR FILING DATE: 1997-11-24  
 ; PRIOR APPLICATION NUMBER: 60/066772  
 ; PRIOR FILING DATE: 1997-11-24  
 ; PRIOR APPLICATION NUMBER: 60/069335  
 ; PRIOR FILING DATE: 1997-12-11  
 ; PRIOR APPLICATION NUMBER: 60/069425  
 ; PRIOR FILING DATE: 1997-12-12  
 ; PRIOR APPLICATION NUMBER: 60/069870  
 ; PRIOR FILING DATE: 1997-12-17  
 ; PRIOR APPLICATION NUMBER: 60/068017  
 ; PRIOR FILING DATE: 1997-12-18  
 ; PRIOR APPLICATION NUMBER: 60/077450  
 ; PRIOR FILING DATE: 1998-03-10  
 ; PRIOR APPLICATION NUMBER: 60/077632  
 ; PRIOR FILING DATE: 1998-03-11  
 ; PRIOR APPLICATION NUMBER: 60/077649  
 ; PRIOR FILING DATE: 1998-03-11  
 ; PRIOR APPLICATION NUMBER: 60/078886  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/078939  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/079664  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/079786  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/080107  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080194  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080327  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/080333  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/081049  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081070  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081195  
 ; PRIOR FILING DATE: 1998-04-09  
 ; PRIOR APPLICATION NUMBER: 60/081838  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/082568  
 ; PRIOR FILING DATE: 1998-04-21  
 ; PRIOR APPLICATION NUMBER: 60/082569

```

RESULT 11
US-09-925-301-1017
; Sequence 1017, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 1017
; LENGTH: 248

```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1017

Query Match      33.4%; Score 408; DB 10; Length 248;
Best Local Similarity 40.2%; Pred. No. 1.1e-29;
Matches 84; Conservative 31; Mismatches 82; Indels 12; Gaps 7;

QY 30 FSRIVGENAVLGSWPMQVSLQDSSGFHFCGSLISQSVVYTAACHNVSV---PGRHFW 86
Db 7 YSRIVGGNMSLLSQWPMQASLQ-FQGYHLCGGSVITPLWIIITAAHCYVDLYLPKSWTIQ 65

QY 87 LGEYDRSSNAEPLQVLSVRAITHPSWNSSTTMNDVTLKLKSPAQYTRISPVCLASN 146
Db 66 VGLVSLLDNPAPSHL--VEKIVYHSKYKPKRLGNDIALMKLAGPLTFNEMIQVCLPNS 123

QY 147 EALTEGLTCVTWGWRL-SGVGNVTPAHQQVALPLVTNOC--ROYWDSSTIDSMICAG 203
Db 124 ENFPDGKVCWTSGWATEDGAGDASPV-LNHAAPVPLISNKHICNHRDVGIIISPSMLCAG 182

QY 204 --GAGASSCGDGGPLVCOKGNTWVLIG 230
Db 183 YLTGGVDSQCGDGGPLVCQERRLKLVG 211

RESULT 12
US-09-804-156-16
; Sequence 16, Application US/09804156
; Patent No. US20020068320A1
; GENERAL INFORMATION:
; APPLICANT: Shi et al.
; TITLE OF INVENTION: Serine protease Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: PT00594
; CURRENT APPLICATION NUMBER: US/09/804,156
; CURRENT FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/189,025
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-804-156-16

Query Match      33.4%; Score 408; DB 10; Length 327;
Best Local Similarity 40.2%; Pred. No. 1.5e-29;
Matches 84; Conservative 31; Mismatches 82; Indels 12; Gaps 7;

QY 30 FSRIVGENAVLGSWPMQVSLQDSSGFHFCGSLISQSVVYTAACHNVSV---PGRHFW 86
Db 86 YSRIVGGNMSLLSQWPMQASLQ-FQGYHLCGGSVITPLWIIITAAHCYVDLYLPKSWTIQ 144

QY 87 LGEYDRSSNAEPLQVLSVRAITHPSWNSSTTMNDVTLKLKSPAQYTRISPVCLASN 146
Db 145 VGLVSLLDNPAPSHL--VEKIVYHSKYKPKRLGNDIALMKLAGPLTFNEMIQVCLPNS 202

QY 147 EALTEGLTCVTWGWRL-SGVGNVTPAHQQVALPLVTNOC--ROYWDSSTIDSMICAG 203
Db 145 VGLVSLLDNPAPSHL--VEKIVYHSKYKPKRLGNDIALMKLAGPLTFNEMIQVCLPNS 202

QY 147 EALTEGLTCVTWGWRL-SGVGNVTPAHQQVALPLVTNOC--ROYWDSSTIDSMICAG 203
Db 203 ENFPDGKVCWTSGWATEDGAGDASPV-LNHAAPVPLISNKHICNHRDVGIIISPSMLCAG 261

QY 204 --GAGASSCGDGGPLVCOKGNTWVLIG 230
Db 262 YLTGGVDSQCGDGGPLVCQERRLKLVG 290

RESULT 13
US-09-946-633-8
; Sequence 8, Application US/09946633
; Patent No. US20020119925A1
; GENERAL INFORMATION:
; APPLICANT: NI et al.
; TITLE OF INVENTION: PT00591
```

```
; FILE REFERENCE: Serine proteases
; CURRENT APPLICATION NUMBER: US/09/946,633
; CURRENT FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 09/597,839
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 60/133,239
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: 60/135,163
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 60/147,005
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: 60/152,935
; PRIOR FILING DATE: 1999-09-09
; PRIOR APPLICATION NUMBER: 60/162,979
; PRIOR FILING DATE: 1999-11-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-946-633-8

Query Match      33.4%; Score 408; DB 10; Length 327;
Best Local Similarity 40.2%; Pred. No. 1.5e-29;
Matches 84; Conservative 31; Mismatches 82; Indels 12; Gaps 7;

QY 30 FSRIVGENAVLGSWPMQVSLQDSSGFHFCGSLISQSVVYTAACHNVSV---PGRHFW 86
Db 86 YSRIVGGNMSLLSQWPMQASLQ-FQGYHLCGGSVITPLWIIITAAHCYVDLYLPKSWTIQ 144

QY 87 LGEYDRSSNAEPLQVLSVRAITHPSWNSSTTMNDVTLKLKSPAQYTRISPVCLASN 146
Db 145 VGLVSLLDNPAPSHL--VEKIVYHSKYKPKRLGNDIALMKLAGPLTFNEMIQVCLPNS 202

QY 147 EALTEGLTCVTWGWRL-SGVGNVTPAHQQVALPLVTNOC--ROYWDSSTIDSMICAG 203
Db 203 ENFPDGKVCWTSGWATEDGAGDASPV-LNHAAPVPLISNKHICNHRDVGIIISPSMLCAG 261

QY 204 --GAGASSCGDGGPLVCOKGNTWVLIG 230
Db 262 YLTGGVDSQCGDGGPLVCQERRLKLVG 290

RESULT 14
US-09-888-615-103
; Sequence 103, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 103
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-103

Query Match      33.4%; Score 408; DB 10; Length 454;
Best Local Similarity 40.2%; Pred. No. 2.3e-29;
Matches 84; Conservative 31; Mismatches 82; Indels 12; Gaps 7;
```

Qy 30 FSQRIVNGENAVLGSWPQVSLQDSSGFHFCGSLISQSWVYVTAHCNVS---FGRHFVV 86  
Db 213 YSSRIVGNGMSLLSQWPMQASLQ-FQGYHLCCGGVITPLWITTAARHCYVDLYLPKSWTIQ 271  
Qy 87 LGEYDRSSNAEPLOVLSVSRATHPSWNSTTMNDVTLKLLASPAQYTTTRISPVCLASSN 146  
Db 272 VGLVSLLDNPAPSHL--VEKIVYHSKYKPKRLGNDIALMKLAGPLTFNEMIQVCLPNS 329  
Qy 147 EALTEGLTCVTTGMRLL-SGVGNVTPAHLQOVALPLVTNOC--ROYWSSITDSMICAG 203  
Db 330 ENFPDGKVCWTSNGATGAGDASPV-LNHAAPLISNKICNHRDVIYGGIISPSMLCAG 388  
Qy 204 --GAGASSCQDGGPLVCQKGNWTWVLIG 230  
Db 389 YLTGVDSCQDGGPLVCQERRLWKLVG 417

RESULT 15  
US-09-967-386-1  
; Sequence 1, Application US/09967386  
; Patent No. US2002015992A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Henkin, Jack  
; APPLICANT: Davidson, Donald J.  
; TITLE OF INVENTION: ANTIANGIOGENIC POLYPEPTIDES AND METHODS  
; FILE REFERENCE: 6738.US.02  
; CURRENT APPLICATION NUMBER: US/09/967,386  
; CURRENT FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: US 60/236,550  
; PRIOR FILING DATE: 2000-09-29  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 791  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-967-386-1

Query Match 32.5%; Score 397; DB 9; Length 791;  
Best Local Similarity 39.3%; Pred. No. 4.6e-28;  
Matches 88; Conservative 35; Mismatches 83; Indels 18; Gaps 8;  
Qy 16 SWGGGIPAIPALSFOSRIVNGENAVLGSWPQVSLQDSSGFHFCGSLISQSWVYVTAAH 75  
Db 545 SFDGKGFQVEPK-KCPGRVGGCVAHPSHPWQVSLRFRGMHFCCGGTLLSPWVLTAAH 603  
Qy 76 C---NVSPGRHFVVLGEYDRSSNAEP-LQVLSVSRATHPSWNSTTMNDVTLKLLASPA 131  
Db 604 CLEKSPRPSSYKVLGAH-QEVNLEPHVQETEVSRLEPT-----RKDIALKLLSSPA 656  
Qy 132 QYTRISPVCLASSNEALTEGLTCVTTGWRGLSGVGNVTPAHLQOVALPLVTNOCROY- 190  
Db 657 VITDKVIPACLPSPNYVADRTECFITGWGETQGTFGA--GLLKEAQLPVIENKVCNRYE 714  
Qy 191 -WSSSITDSMICAG--GAGASSCQDGGPLVCQKGNWTWVLIG 231  
Db 715 FLNGRVSTELCAGHLAGGTDCQDGGPLVCFCFKDKYILQGV 758

Search completed: December 20, 2002, 15:21:06  
Job time : 8.30435 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:14:08 ; Search time 28.6087 Seconds  
(without alignments)  
1663.721 Million cell updates/sec

Title: US-09-856-319B-2\_COPY\_1\_231

Perfect score: 1221

Sequence: 1 MLLSLTLTLVLLSSWGCG.....GDSGGPLVCQKGNWTWVLGI 231

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rviris.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1102	90.3	264	11 Q9EQZ8	Q9eqz8 rattus norv
2	1093	89.5	264	11 Q9D7P8	Q9d7p8 mus musculu
3	1093	89.5	264	11 Q9ER05	Q9er05 mus musculu
4	1083	88.7	264	11 Q9D960	Q9d960 mus musculu
5	766.5	62.8	261	13 Q9W7Q4	Q9w7q4 paralichthy
6	710.5	58.2	260	13 Q9W7Q3	Q9w7q3 paralichthy
7	684.5	56.1	263	11 Q9DC86	Q9dc86 mus musculu
8	682.5	55.9	263	11 Q9CR35	Q9cr35 mus musculu
9	676.5	55.4	263	11 Q9D8X8	Q9d8x8 mus musculu
10	674.5	55.2	263	13 Q9PWQ6	Q9pwq6 gadus morhu
11	472.5	38.7	269	11 Q9C052	Q9cq52 mus musculu
12	465.5	38.1	266	13 Q92077	Q92077 gadus morhu
13	464.5	38.0	269	11 Q9D7T9	Q9d7t9 mus musculu
14	454.5	37.2	260	13 Q9W7P9	Q9w7p9 paralichthy
15	447	36.6	164	11 Q9DC82	Q9dc82 mus musculu
16	442	36.2	249	13 Q9W7Q1	Q9w7q1 paralichthy

17	434	35.5	269	4 Q96QV5	Q96qv5 homo sapien
18	423	34.6	267	5 Q9BK47	Q9bk47 luidia foli
19	422	34.6	270	13 Q91039	Q91039 gadus morhu
20	420.5	34.4	270	4 Q96QL8	Q96ql8 homo sapien
21	420	34.4	277	5 Q96899	Q96899 scolopendra
22	417.5	34.2	266	13 Q9W7Q0	Q9w7q0 paralichthy
23	417.5	34.2	269	6 Q95KW7	Q95kw7 bos taurus
24	411	33.7	268	13 Q9W7Q2	Q9w7q2 paralichthy
25	409	33.5	1524	13 Q91674	Q91674 xenopus lae
26	400.5	32.8	257	6 Q19023	Q19023 macaca mula
27	400.5	32.8	339	11 Q99144	Q99144 mus musculu
28	399.5	32.7	453	11 Q8VDE0	Q8vde0 mus musculu
29	399	32.7	263	5 Q62562	Q62562 penaeus van
30	398.5	32.6	264	13 Q8QGF6	Q8qgf6 xenopus lae
31	398	32.6	810	4 Q15146	Q15146 homo sapien
32	392.5	32.1	812	11 Q91WJ5	Q91wj5 mus musculu
33	392	32.1	334	6 Q46507	Q46507 papio hamad
34	389.5	31.9	247	13 Q42608	Q42608 petromyzon
35	389.5	31.9	767	13 Q9DGR2	Q9dgr2 xenopus lae
36	388.5	31.8	247	13 Q42158	Q42158 petromyzon
37	388.5	31.8	331	11 Q8R1A6	Q8rla6 mus musculu
38	386	31.6	787	5 Q9VEY6	Q9vey6 drosophila
39	386	31.6	1004	13 P79953	P79953 xenopus lae
40	385.5	31.6	273	11 Q921N4	Q921n4 mus musculu
41	385.5	31.6	812	11 Q9ROW3	Q9row3 rattus norv
42	381	31.2	266	6 Q46644	Q46644 macaca fasc
43	381	31.2	266	11 Q91X79	Q91x79 mus musculu
44	380	31.1	454	6 Q46506	Q46506 papio hamad
45	379	31.0	492	4 Q96T73	Q96t73 homo sapien

## ALIGNMENTS

RESULT 1

Q9EQZ8 ID Q9EQZ8 PRELIMINARY; PRT; 264 AA.  
AC Q9EQZ8;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Chymopasin.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=RAT PANCREAS;  
RA Sogame Y., Mitsui S., Kataoka K., Kashima K., Kato M., Sakagami J.,  
RA Yanaguchi N.;  
RT "Molecular cloning of rat chymopasin.";  
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
CC TRYPSIN FAMILY.  
DR EMBL; AB020757; BAB20287.1; -;  
DR HSSP; P00766; 4CHA.  
DR MEROPS; S01.256; -;  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR001254; Ser\_protease\_Try.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; Tryp\_Spc; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease.  
SQ SEQUENCE 264 AA; 28116 MW; F9FD5D210FD3500E CRC64;

Query Match 90.3%; Score 1102; DB 11; Length 264;  
Best Local Similarity 87.0%; Pred. No. 5,1e-92;  
Matches 201; Conservative 17; Mismatches 13; Indels 0; Gaps 0;

```

QY 1 MLLSLTSLVLGSGWCGGIPAIPALSFORSORIVNGENAVLGSWPQVSLQDSSGFHFC 60
Db 1 MLLSLTSLVLGSGWCGGIPAIPALSFORSORIVNGENAVLGSWPQVSLQDSSGFHFC 60
QY 61 GGSLSQSWVVTAAHCNVSPGRHFVVLGEYDRSSNAEPQLVLSVRAITHPSWNSTMMN 120
Db 61 GGSLSQSWVVTAAHCNVSPGRHFVVLGEYDRSSNAEPQLVLSVRAITHPSWNSTMMN 120
QY 121 DVTLLKASPAQYTTTRISPVCLASSNEALTEGLTCVTTGWRSLSGVGNVTPAHLQOVALP 180
Db 121 DVTLLKASPAQYTTTRISPVCLASSNEALTEGLTCVTTGWRSLSGVGNVTPAHLQOVALP 180
QY 181 LVTVNOCROYWDSITDSMICAGGASSCGQSGGGLVCQKGTWVLIGI 231
Db 181 LVTVNOCROYWDSITDSMICAGGASSCGQSGGGLVCQKGTWVLIGI 231

RESULT 2
Q9D7P8 PRELIMINARY; PRT; 264 AA.
AC Q9D7P8;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 1810004D15rik protein.
GN CTRL OR 1810004D15RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Cariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT Functional annotation of a full-length mouse cDNA collection.;
RL Nature 409:685-690(2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AK009019; BAB26029.1; -.
DR HSSP; P00766; 4CHA.
DR MEROPS; S01.256; -.
DR MGD; MGI:88558; Ctrl.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 264 AA; 28167 MW; 1D979469A07056C2 CRC64;

Query Match 89.5%; Score 1093; DB 11; Length 264;
Best Local Similarity 85.7%; Pred. No. 3.3e-91;
Matches 198; Conservative 21; Mismatches 12; Indels 0; Gaps 0;

```

```

Matches 198; Conservative 21; Mismatches 12; Indels 0; Gaps 0;
QY 1 MLLSLTSLVLGSGWCGGIPAIPALSFORSORIVNGENAVLGSWPQVSLQDSSGFHFC 60
Db 1 MLLSLTSLVLGSGWCGGIPAIPALSFORSORIVNGENAVLGSWPQVSLQDSSGFHFC 60
QY 61 GGSLSQSWVVTAAHCNVSPGRHFVVLGEYDRSSNAEPQLVLSVRAITHPSWNSTMMN 120
Db 61 GGSLSQSWVVTAAHCNVSPGRHFVVLGEYDRSSNAEPQLVLSVRAITHPSWNSTMMN 120
QY 121 DVTLLKASPAQYTTTRISPVCLASSNEALTEGLTCVTTGWRSLSGVGNVTPAHLQOVALP 180
Db 121 DVTLLKASPAQYTTTRISPVCLASSNEALTEGLTCVTTGWRSLSGVGNVTPAHLQOVALP 180
QY 181 LVTVNOCROYWDSITDSMICAGGASSCGQSGGGLVCQKGTWVLIGI 231
Db 181 LVTVNOCROYWDSITDSMICAGGASSCGQSGGGLVCQKGTWVLIGI 231

RESULT 3
Q9ER05 PRELIMINARY; PRT; 264 AA.
AC Q9ER05;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Chymopasin (Chymotrypsin A CTRA-1).
GN CTRL OR CTRA1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Mitsui S., Yamaguchi N.;
RT "Molecular cloning of mouse chymopasin.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129S6/SVEVTC; TISSUE=SPLEEN;
RA Bjoernslett M.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AB016228; BAB20275.1; -.
DR EMBL; AF236365; AAL11034.1; -.
DR HSSP; P00766; 4CHA.
DR MEROPS; S01.256; -.
DR MGD; MGI:88558; Ctrl.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 264 AA; 28135 MW; 1D979709A07056C2 CRC64;

Query Match 89.5%; Score 1093; DB 11; Length 264;
Best Local Similarity 85.7%; Pred. No. 3.3e-91;
Matches 198; Conservative 21; Mismatches 12; Indels 0; Gaps 0;
QY 1 MLLSLTSLVLGSGWCGGIPAIPALSFORSORIVNGENAVLGSWPQVSLQDSSGFHFC 60
Db 1 MLLSLTSLVLGSGWCGGIPAIPALSFORSORIVNGENAVLGSWPQVSLQDSSGFHFC 60
QY 61 GGSLSQSWVVTAAHCNVSPGRHFVVLGEYDRSSNAEPQLVLSVRAITHPSWNSTMMN 120
Db 61 GGSLSQSWVVTAAHCNVSPGRHFVVLGEYDRSSNAEPQLVLSVRAITHPSWNSTMMN 120
QY 121 DVTLLKASPAQYTTTRISPVCLASSNEALTEGLTCVTTGWRSLSGVGNVTPAHLQOVALP 180
Db 121 DVTLLKASPAQYTTTRISPVCLASSNEALTEGLTCVTTGWRSLSGVGNVTPAHLQOVALP 180
QY 181 LVTVNOCROYWDSITDSMICAGGASSCGQSGGGLVCQKGTWVLIGI 231
Db 181 LVTVNOCROYWDSITDSMICAGGASSCGQSGGGLVCQKGTWVLIGI 231

Query Match 89.5%; Score 1093; DB 11; Length 264;
Best Local Similarity 85.7%; Pred. No. 3.3e-91;
Matches 198; Conservative 21; Mismatches 12; Indels 0; Gaps 0;

```

Db	121	DLTLKLASPARYAQSPVCLASTNEALPUGTCTVTTGWRISGVNTPARLQOVL	180
Qy	181	LVTVNCRQYWDSSITDSMICAGAGASSCGQSDGGPLVCQKGNWTWLIGI	231
Db	181	LVTVNCRQYWGARIITAMICAGGAGASSCGQSDGGPLVCQKGNWTWLIGI	231
RESULT 4			
Q9D960			
ID	Q9D960	PRELIMINARY; PRT; 264 AA.	
AC	Q9D960;		
DT	01-JUN-2001 (TrEMBLrel. 17, Created)		
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)		
DE	1810004D15Rik protein.		
GN	1810004D15Rik		
DE	Ctrl OR 1810004D15RIK.		
OS	Mus musculus (Mouse)		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAN=C57BL/6J; TISSUE=PANCREAS;		
RC	MEDLINE=21085660; PubMed=11217851;		
RA	Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,		
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,		
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,		
RA	Kadoya K., Matsuda H.A., Asnburner M., Batalov S., Casavant T.,		
RA	Fleischmann W., Gaasterland T., Glassi C., King B., Kochiwa H.,		
RA	Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,		
RA	Schral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,		
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,		
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,		
RA	Brustein M.J., Bult C., Fietcher C., Fujita M., Gariboldi M.,		
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,		
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,		
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,		
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,		
RA	Suzuki H., Toyo-oka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,		
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kotsuki S.,		
RA	Hayashizaki Y.;		
RT	Functional annotation of a full-length mouse cdna collection.;		
RL	Nature 409:685-690(2001).		
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.		
DR	EMBL; AK007333; BAB24967.1; -.		
DR	HSSP; P00766; 4CHA.		
DR	MEROPS; S01.256; -.		
DR	MGI; MGI:88558; Ctrl.		
DR	InterPro; IPR001314; Chymotrypsin.		
DR	InterPro; IPR001254; Ser-protease_Try.		
DR	Pfam; PF00089; trypsin.1.		
DR	PRINTS; PR00722; CHYMOTRYPSIN.		
DR	SMART; SM00020; Tryp_Spc.1.		
DR	PROSITE; PS50240; TRYPSIN_DOM; 1.		
DR	PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.		
DR	PROSITE; PS00135; TRYPSIN_SER; 1.		
DR	Hydrolase; Serine protease.		
QW	SEQUENCE 264 AA; 28151 MW; 1D979719E07C16DE CRC64;		

	Query Match	88.7%	Score 1083;	DB 11;	Length 264;
	Best Local Similarity	85.3%;	Pred. No. 2.7e-90;		
	Matches 197;	Conservative 21;	Mismatches 13;	Indels 0;	Gaps 0;
Qy	1	MLLSLTLSLVLLGSSNGCGIPATKPAISPSQRTVINGENAVLGSWPQVSLQDSSGGHFHC	60		
Db	1	MLLSLTLTLVLLGSSNGCGIPATPALSYNQRTVINGENAVPGSWPQVSLQDNTGPFHC	60		
Qy	61	GGSLISGSWVYTAACHNVSFCGRHFVVLGEYDRSSNAPQLQVLSVRAITHPSWNSITTMNN	120		
Db	61	GGSLISPNWVYTAACHQVTPGRHFVWLGEYDRSSNAPQVQLSTARITHPNWNAITMNN	120		

Qy	121	DVTLKLLASPAQYTRTRISPCVLASSNEALTEGLTCTVTGWRGSLSGVGNVTPAHLLQQVALP	180
Db	121	DLTLKLLASPARTAQYVCLASTNEALPSGLTCTVTGMRISGSGVGNVTPARLQQVLP	180
Qy	181	LVTVNCRQYWDSSITDSMTICAGGAGASSCGDGGGGLVCQKNTWYLYGI	231
Db	181	LVTVNCRQYWGARIITDAMTCAGSGASSCGDGGGGLVCQKNTWYLYGI	231

```

RESULT 5
Q9W7Q4
ID Q9W7Q4 PRELIMINARY; PRT; 261 AA.
AC Q9W7Q4;
DT 01-NOV-1999 (T:EMBLrel. 12, Created)
DT 01-NOV-1999 (T:EMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (T:EMBLrel. 21, Last annotation update)
DE Chymotrypsinogen 1.
OS Paralicthys olivaceus (Flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidae; Paralichthyidae; Paralichthys.
NCBI_TaxID=8255;
RX [1]
RN
RX
RN
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RC Suzuki T., Srivastava A.S., Kurokawa T.;
RT "Japanese flounder mRNA for chymotrypsinogen 1.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AB029753; BAA82365.1; -.
DR HSSP; P00763; IDPO.
DR MCRPS; S01.256; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Hydrolase. Serine protease.
SQ SEQUENCE 261 AA; 28184 MW. D7090A06539587D CRC64;

```

Query Match	62.8%	Score	766.5	DB	13	Length	261
Best Local Similarity	61.6%	Pred. No.	1.3e-61				
Matches	143	Conservative	31	Mismatches	53	Indels	5
							Gaps 3
Qy	1	MLLSLTLSLVLLGSSWCCGIPAKLPALSFQRIVNGENAVLGSWPQVSLQDSSGFHFC	60				
Db	1	MLLFISFCALV--ASALGCGVPSIKPQSGVYKNIKNGETAIVSGSWPQVSLQDGRGFHFC	58				
Qy	61	GGSLISGSWVYTAACHCNVSPGRHFVWLCEYDRSSNAPQLQVLSVSRATHTPSMNSTTMNN	120				
Db	59	GGSLISYVWVYTAACHVSPNRHRVILGEHDROYNEPIQVMSATARAHTPHYNSQNFNN	118				
Qy	121	DVTLLKLASPAQYTRISPYCVCLASNEALTEGLTCVTTGWGRLSVGVNVTPAHLLQQVALP	180				
Db	119	DITLLRLSSPQMTSRVSPVCLASSSTSIPTGTCVTTGWGRTGQTS--SPRYLQQTSLP	176				
Qy	181	LVTYNOCROYWD--SITDSMTICAGACASSCGDGGPLVCOKGNVAVLIGI	231				
Db	177	LLSPAQCKQYGYNRIITDAMICAGASVSSCGDGGPLVCEKSGAMFLTGI	228				
RESULT	6						
Q9W7Q3		PRELIMINARY		PRT		260	AA
ID	Q9W7Q3						
AC	Q9W7Q3						
DT	01-NOV-1999	(TrEMBLrel. 12, Created)					
DT	01-NOV-1999	(TrEMBLrel. 12, Last sequence update)					

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Chymotrypsinogen 2.  
OS Paralicthys olivaceus (Flounder).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;  
OC Pleuronectoidae; Paralichthyidae; Paralichthys.  
OX NCBI\_TaxID=8255;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PANCREAS;  
RA Suzuki T., Srivastava A.S., Kurokawa T.;  
RT "Japanese flounder mRNA for chymotrypsinogen 2.";  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
CC TRYPSIN FAMILY.  
DR EMBL: AB029754; BAA82366.1; -;  
DR HSSP: P00766; LCHG.  
DR MEROPS: S01.152; -;  
DR InterPro: IPR001314; Chymotrypsin.  
DR InterPro: IPR001254; Ser\_protease\_Try.  
DR Pfam: PF00089; trypsin; 1.  
DR PRINTS: PR00722; CHYMOTRYPSIN.  
DR SMART: SM00020; Tryp\_Spc; 1.  
DR PROSITE: PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE: PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease.  
SQ SEQUENCE 260 AA; 27793 MW; 9F583044E22F78C0 CRC64;

Query Match 58.2%; Score 710.5; DB 13; Length 260;  
Best Local Similarity 57.7%; Pred. No. 1.6e-56;  
Matches 128; Conservative 36; Mismatches 57; Indels 1; Gaps 1;  
QY 10 LVLLGSGWGGIPAIKPAISFQRIYNGENAVLGSWPMQVSLQDSSGFHFCGSGSLISQSW 69  
DB 7 LAFAGAAICGSGIPAIPIVGYIRYNGEALPHSPWQVSLQDYTGFFHCGGSLINEW 66  
QY 70 VVTAHCNVSFGRHVVYLGEDRSSNAEPLQVLSVRAITHPSWNTTNNVTLKLKAS 129  
DB 67 VVTAHCNVRTS-HRVILGHDHSSAEDIQVMKVKFKPRYNGTYTINNDILLIKLAA 125  
QY 130 PAQYTRISPVCLASNEALTEGLTCVTTGWRSLGSGVNTVPAHLOQVALPLVTYNQCRQ 189  
DB 126 PAQNMNRVSPVCAETSDNFPAGMKCVTSGWGLTRHNPDTPALLOQAAALPLTNDCCR 185  
QY 190 YWDSSTITDSMICAGGAGASCQDGGPLVCQKGNWVLIGI 231  
DB 186 YWGNKISLNMICAGAGASCQDGGPLVCQKAGATVLGI 227

RESULT 7  
Q9DC86 ID Q9DC86 PRELIMINARY; PRT; 263 AA.  
AC Q9DC86  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE 220000809Rik protein.  
GN 220000809Rik.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=SPLEEN;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Futuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann J., Mazzarelli J., Monbaerts P.,  
RA Lyons P., Marchionni L., Mashima J., Mazzaresli J., Momamoto N.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
RA Hayashizaki Y.;  
RA "Functional annotation of a full-length mouse cDNA collection.";  
RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
RL Nature 403:685-690(2001).  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
CC TRYPSIN FAMILY.  
DR EMBL: AK003060; BAB22539.1; -;  
DR HSSP: P00766; LGCT.  
DR MEROPS: S01.152; -;  
DR MGD: MGI:1913723; 220000809Rik.  
DR InterPro: IPR001314; Chymotrypsin.  
DR InterPro: IPR001254; Ser\_protease\_Try.  
DR Pfam: PF00089; trypsin; 1.  
DR PRINTS: PR00722; CHYMOTRYPSIN.  
DR SMART: SM00020; Tryp\_Spc; 1.  
DR PROSITE: PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE: PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease.  
SQ SEQUENCE 263 AA; 27821 MW; 2620A27AFBA5D04D CRC64;

Query Match 56.1%; Score 684.5; DB 11; Length 263;  
Best Local Similarity 54.5%; Pred. No. 3.7e-54;  
Matches 126; Conservative 38; Mismatches 66; Indels 1; Gaps 1;  
QY 1 MLLLSITLVLGSGWGGIPAIKPAISFQRIYNGENAVLGSWPMQVSLQDSSGFHFC 60  
DB 1 MAPFLVSCFALVGFPGVPAIQVLTGLSRVNGEDAIPTGSPWQVSLQDRTGFHFC 60  
QY 61 GGLISQSWVTAHCNVSFGRHVVYLGEDRSSNAEPLQVLSVRAITHPSWNTTNN 120  
DB 61 GGLISENVVTAHCNVSFGRHVVYLGEDRSSNAEPLQVLSVRAITHPSWNTTNN 119  
QY 121 DVTLLKASPAQYTRISPVCLASNEALTEGLTCVTTGWRSLGSGVNTVPAHLOQVALP 180  
DB 120 DITLLKATPAQFSETVSAVCLPTVDDFFPAGTLCATTGKTKYNALKTDPKLOQAALP 179  
QY 181 LVTVNCROYWDSSTITDSMICAGGAGASCQDGGPLVCQKGNWVLIGI 231  
DB 180 IVSEAKCKESWGSKITDVMICAGAGVSSCMGDSGGLPLVCQKQGVWTLGI 230

RESULT 8  
Q9CR35 ID Q9CR35 PRELIMINARY; PRT; 263 AA.  
AC Q9CR35  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE 220000809Rik protein.  
GN 220000809Rik.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=STOMACH, SPLEEN, AND PANCREAS;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

```
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-Oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC EMBL: AK008927; BAB25971.1; -.
DR EMBL: AK003079; BAB22553.1; -.
DR EMBL: AK007765; BAB25241.1; -.
DR EMBL: AK007815; BAB25280.1; -.
DR EMBL: AK008729; BAB25861.1; -.
DR EMBL: AK008888; BAB25954.1; -.
DR HSSP: P00766; IGCT.
DR MEROPS: S01.152; -.
DR MGD: MGI:1913723; 2200008D09Rik.
DR InterPro: IPR001314; Chymotrypsin.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; Tryp_SPC; 1.
DR PROSITE: PS02040; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 263 AA; 27822 MW; 28C4487AF1A26B27 CRC64;

Query Match 55.9%; Score 682.5; DB 11; Length 263;
Best Local Similarity 54.5%; Pred. No. 5.6e-54;
Matches 126; Conservative 37; Mismatches 57; Indels 1; Gaps 1;

QY 1 MLLSLTSLVLGSSWGGCIPAIKPAISQRIYNGENAVLGSWPQVSLQDSSGFHFC 60
Db 1 MAFLWLVSFCALVGAFEGCGVPAIQVLTGLSRIVNGEDAIPGSWPQVSLQDRTGFHFC 60
QY 61 GGSLSIQSWVTAHNCVSPGRHFVVLGVEYDRSSNAEPLQVLSVSRATHPSWNSTTMN 120
Db 61 GGSLSIENWVTAHNCVSPGRHFVVLGVEYDRSSNAEPLQVLSVSRATHPSWNSTTMN 119
QY 121 DVTLLKASPAQYTRTRISPVCLASNEALTEGLTCVTTGWRGLSGVNTVPAHLOQVALP 180
Db 120 DITLLKLATPAQFSETVSAVCLPTVDDPDPAGTLCATGWTGKTKYNALKTDPKLOQAALP 179
QY 181 LVTYNOCROYWDSSTIDSMICAGGASCCGDSGGPLVCQKGNVTWVLGI 231
Db 180 IVSEAKCKESWGSKITDVMICAGSGVSSCMGDSGGPLVCQKQDGVWTLGI 230

RESULT 9
Q9D8X8 PRELIMINARY; PRT; 263 AA.
AC Q9D8X8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 2200008D09Rik protein.
GN Mus musculus (Mouse).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
```

```
SEQUENCE FROM N.A.
STRAIN=C57BL/6J; TISSUE=PANCREAS;
MEDLINE=21085660; PubMed=11217851;
Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
Suzuki H., Toyo-Oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC EMBL: AK007566; BAB25112.1; -.
DR HSSP: P00766; IGCT.
DR MEROPS: S01.152; -.
DR MGD: MGI:1913723; 2200008D09Rik.
DR InterPro: IPR001314; Chymotrypsin.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; Tryp_SPC; 1.
DR PROSITE: PS02040; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 263 AA; 27898 MW; C0638FB8F905A92F CRC64;

Query Match 55.4%; Score 676.5; DB 11; Length 263;
Best Local Similarity 54.1%; Pred. No. 1.9e-53;
Matches 125; Conservative 37; Mismatches 68; Indels 1; Gaps 1;

QY 1 MLLSLTSLVLGSSWGGCIPAIKPAISQRIYNGENAVLGSWPQVSLQDSSGFHFC 60
Db 1 MAFLWLVSFCALVGAFEGCGVPAIQVLTGLSRIVNGEDAIPGSWPQVSLQDRTGFHFC 60
QY 61 GGSLSIQSWVTAHNCVSPGRHFVVLGVEYDRSSNAEPLQVLSVSRATHPSWNSTTMN 120
Db 61 GGSLSIENWVTAHNCVSPGRHFVVLGVEYDRSSNAEPLQVLSVSRATHPSWNSTTMN 119
QY 121 DVTLLKASPAQYTRTRISPVCLASNEALTEGLTCVTTGWRGLSGVNTVPAHLOQVALP 180
Db 120 DITLLKLATPAQFSETVSAVCLPTVDDPDPAGTLCATGWTGKTKYNALKTDPKLOQAALP 179
QY 181 LVTYNOCROYWDSSTIDSMICAGGASCCGDSGGPLVCQKGNVTWVLGI 231
Db 180 IVSEAKCKESWGSKITDVMICAGSGVSSCMGDSGGPLVCQKQDGVWTLGI 230

RESULT 10
Q9PWQ6 PRELIMINARY; PRT; 263 AA.
AC Q9PWQ6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 2200008D09Rik protein.
GN Chymotrypsin B precursor (EC 3.4.21.1).
OS Gadus morhua (Atlantic cod).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.  
 OX NCBI\_TaxID=8049;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PYLORIC CAECA;  
 RX MEDLINE=20464334; PubMed=11011764;  
 RA Spilliaert R., Gudmundsdottir A.;  
 RT "Molecular Cloning of the Atlantic Cod Chymotrypsinogen B.";  
 RL Microb. Comp. Genomics 5:41-50(2000).  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPSIN FAMILY.  
 DR EMBL; AJ242521; CAB43766.1; -.  
 DR HSP; P00766; 1CHG.  
 DR MEROPS; S01.152; -.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Hydrolase; Serine protease; Signal.  
 FT SIGNAL 1 16 POTENTIAL.  
 FT CHAIN 17 263 CHYMOTRYPSIN B.  
 SQ SEQUENCE 263 AA; EF61B18A34EE5E7C CRC64;  
 Query Match 55.2%; Score 674.5; DB 13; Length 263;  
 Best Local Similarity 57.0%; Pred. No. 3e-53;  
 Matches 127; Conservative 38; Mismatches 51; Indels 7; Gaps 5;  
 QY 13 LGSSWGGGIPAIKPAIFSQRVINGENAVLGSPWQVSLQDSSGFHPCGSLISQSMVVT 72  
 DB 11 ISAIYGGSPAIQPVQYGIARVINGEAVPHSPWQVSLQDSNGFHCGLSLINENWVVT 70  
 QY 73 AAHCNVSPGRHVFVVLGVDYRS-SNAEPLOVLSVRAITHPSWNTTNNNDVTLKLKASPA 131  
 DB 71 AAHCNVRT-VHRVIVGEHDKSRASDENIQLKPSWVTFPKWDSRTINNDISLIKASPA 129  
 QY 132 QYTRISPVCLASNEALTEGLCTVTGWG--RLSGVGNVTPAHLQVALPLTVNQCRO 189  
 DB 130 VLGNVSPVCLGESSDVFAPGMKCVTSCWGLTRYNAPG--TPNKLOQAALPLMSNEECQ 187  
 QY 190 YW-DSSITDSMICAGGASCGDGGPLVCQKGNWTLIGI 231  
 DB 188 TWGNMISDVNICAGAGATSCWGDGSGPLVCQKDNWTLIGI 230  
 RESULT 11  
 Q9CQ52 ID Q9CQ52 PRELIMINARY; PRT; 269 AA.  
 AC Q9CQ52  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE 2310074F01R4K protein.  
 GN ELA3B OR 2310074F01R4K.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=TONGUE;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shingagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamlya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPSIN FAMILY.  
 DR EMBL; AK010149; BAB26734.1; -.  
 DR EMBL; AK009129; BAB26092.1; -.  
 DR HSP; P05805; IFON.  
 DR MEROPS; S01.154; -.  
 DR MGD; MGI:1915118; ELA3b.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Hydrolase; Serine protease.  
 SQ SEQUENCE 269 AA; 28904 MW; C543F76957B2A7CE CRC64;  
 Query Match 38.7%; Score 472.5; DB 11; Length 269;  
 Best Local Similarity 43.5%; Pred. No. 6.3e-35;  
 Matches 104; Conservative 31; Mismatches 89; Indels 15; Gaps 7;  
 QY 3 LLSLTLSVLGSSWGGIPAIKPAIFSQRVINGENAVLGSPWQVSLQ---DSSGFHF 59  
 DB 1 MLRLSSLLVALASGCGQSPHNP---SSRVNGEAVPHSPWQVSLQYKDGSPHHT 56  
 QY 60 CGSGLISQSMVVTAAHCNVSPGRHVFVVLGVDYRS-SNAEPLOVLSVRA--ITHPSWNS-- 115  
 DB 57 CGSLITPDWVLVAGHCISTRTYQVVLGHERGVEGQGVIPINAGDLFVHPKWNMSMC 116  
 QY 116 TTNMNDVTLKLKASPAQYTRISPVCLASNEALTEGLCTVTGWRLSGVGNVTPAHLQ 175  
 DB 117 VSGNDIALVKLSRQAGDAVQLACLPAGEILPNGAPCYISGWRLSTNGPL-PDKLQ 175  
 QY 176 QVALPLTVNQCRO--YWDSSITDSMICAGGASCGDGGPLVCQKGN-TWVLIGI 231  
 DB 176 QALLPVVDYEHCSRWNWGLSVKTTVMVACGDIQSGCGDGGPLNCPADNGTWQVHGV 234  
 RESULT 12  
 Q92077 ID Q92077 PRELIMINARY; PRT; 266 AA.  
 AC Q92077  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Elastase precursor (EC 3.4.21.37) (Leukocyte elastase) (Lysosomal  
 DE elastase) (Neutrophil elastase) (Bone marrow serine protease)  
 DE (MEDULLASIN).  
 GN ELB.  
 OS Gadus morhua (Atlantic cod).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.  
 OX NCBI\_TaxID=8049;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96262679; PubMed=8925447;  
 RA Gudmundsdottir E., Spilliaert R., Yang Q., Craik C.S., Bjarnason J.B.,  
 RA Gudmundsdottir A.;





DR InterPro; IPR001254; Ser\_protease\_Try.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; TRYP\_SPC; 1.  
DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease.  
FT NON\_TER 1  
SQ SEQUENCE 260 AA; 27942 MW; 61883CD1B6A66075 CRC64;

Query Match 37.2%; Score 454.5; DB 13; Length 260;  
Best Local Similarity 40.3%; Pred. No. 2.6e-33;  
Matches 93; Conservative 45; Mismatches 80; Indels 13; Gaps 8;

QY 8 LSLVLGSSWGGIPALPALESQRIYNGENAVLGSMPQVSLQDSSG--FHFCGSL 64  
DB 1 LALFVAG-AYGGLPTFPPTIT---RVGGDDVRENSWPQVSLQKSGSNFYHTCGGFL 56  
QY 65 ISQWVVTAAHCNVSPGRHFVVLGEYDRSSNAEPLQVLVSRAITHPSWNSTTMNNDVTL 124  
DB 57 ISNQWVLTAAHC-IGSRTRYVYMGKHLRNDKAGSTAISPAKLIIVHNWDSYRINDIAL 115  
QY 125 LKLASPAQYTTTRISPVCLASNEALTEGLTCVTGWRGSLGVGNVTPPAHLQOVALPLVTV 184  
DB 116 IKLATPVTESDSVMAACPLSDGIQPNAGPCVYVTGWRGLWTGGPIADI-LQOALLPVVDH 174  
QY 185 NCRQ--YWDSSITDSMICAGAG-ASSCGSGGGLVCQK-GNTWVLGI 231  
DB 175 ATCTSDMWGSLVNTSMVCAGAGDLASCGDSGGPLNCQSPDGSEVHGCV 225

## RESULT 15

Q9DC82 PRELIMINARY; PRT; 164 AA.  
AC Q9DC82;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE O910001G08Rik protein.  
GN CTRL OR 0910001G08Rik.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-C57BL/6J; TISSUE-SPLEEN;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Salto R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
DR EMBL; AK003074; BAB22549.1; -  
DR HSSP; P00763; IDPO.  
DR MEROPS; S01.997; -.  
DR MGD; MGI:88558; Ctrl.

DR InterPro; IPR001254; Ser\_protease\_Try.  
DR Pfam; PF00089; trypsin; 1.  
DR SMART; SM00020; TRYP\_SPC; 1.  
DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease.  
SQ SEQUENCE 164 AA; 17707 MW; 83791FD829ABEDD6 CRC64;  
Query Match 36.6%; Score 447; DB 11; Length 164;  
Best Local Similarity 69.4%; Pred. No. 7e-33;  
Matches 86; Conservative 9; Mismatches 27; Indels 2; Gaps 2;  
QY 110 HPSWNST-TMNDVTLTKLASP-AQYTTTRISPVCLASNEALTEGLTCVTGWRGSLGVG 167  
DB 8 HPHWGGPHPLKTKTRLFLEFFFPPEKTKFPFGLASTNEALPSGLTCVTGWRISGVG 67  
QY 168 NYTPAHLQOVALPLTVNOCROYWDSSTDSMICAGAGASSCGSGGLVCQKGNWTV 227  
DB 68 NYTPARLQOVLPLTVNOCROYWGARIITDAMICAGGSGASSCGSGGLVCQKGNWTV 127  
QY 228 LIGI 231  
DB 128 LIGI 131

Search completed: December 20, 2002, 15:18:53  
Job time : 29.6087 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:12:58 ; Search time 7.91304 Seconds  
(without alignments)  
1210.790 Million cell updates/sec

Title: US-09-856-319b-2\_copy\_1\_231

Perfect score: 1221

Sequence: 1 MLLSLTSLVLLGSGWCG.....GDSGGPLVCQKGTWVLIGI 231

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query Match %	Score	Length	DB ID	Description
1	1214	99.4	264	1	CTRL_HUMAN
2	714.5	58.5	263	1	CTRL_HUMAN
3	701.5	57.5	263	1	CTRL_HUMAN
4	689.5	56.5	263	1	CTRL_HUMAN
5	670.5	54.9	263	1	CTRL_HUMAN
6	666.5	54.6	245	1	CTRL_HUMAN
7	684.5	54.4	245	1	CTRL_HUMAN
8	657.5	53.8	245	1	CTRL_HUMAN
9	472	38.7	269	1	EL2_PIG
10	453.5	37.1	270	1	EL3B_HUMAN
11	453	37.1	269	1	EL2_BOVIN
12	445	36.4	269	1	EL2A_HUMAN
13	438	35.9	271	1	EL2_RAT
14	427	35.0	269	1	EL2B_HUMAN
15	420.5	34.4	270	1	EL3A_HUMAN
16	415	34.0	271	1	EL2_MOUSE
17	410.5	33.6	253	1	CAC3_BOVIN
18	410.5	33.6	253	1	PSS8_HUMAN
19	408	33.4	454	1	TMS2_HUMAN
20	407.5	33.4	268	1	CLCR_RAT
21	404	33.1	343	1	PLMN_SHEEP
22	398.5	32.6	342	1	PSS8_RAT
23	397	32.5	810	1	PLMN_HUMAN
24	396.5	32.5	342	1	PSS8_MOUSE
25	396	32.4	810	1	PLMN_MACMU
26	395.5	32.4	1035	1	ENTK_BOVIN
27	392	32.1	266	1	EL1_BOVIN
28	391	32.0	790	1	PLMN_PIG
29	390.5	32.0	268	1	CLCR_HUMAN
30	389.5	31.9	812	1	PLMN_MOUSE
31	389	31.9	338	1	PLMN_HORSE
32	388	31.8	270	1	TRYT_MERON
33	386	31.6	786	1	STUB_DROME

34	384.5	31.5	238	1	TRY3_SALSA
35	384.5	31.5	273	1	MCT7_MOUSE
36	383.5	31.4	246	1	TRY1_CANFA
37	383.5	31.4	246	1	TRY1_RAT
38	383.5	31.4	275	1	TRYT_PIG
39	382.5	31.3	1069	1	ENTK_MOUSE
40	382	31.3	248	1	TRY1_CHICK
41	382	31.3	248	1	TRY2_CHICK
42	382	31.3	1019	1	ENTK_HUMAN
43	380	31.1	333	1	PLMN_CANFA
44	379	31.0	248	1	TRY3_CHICK
45	379	31.0	492	1	TMS2_HUMAN

P35033 salmo salar  
Q02844 mus musculus  
P06871 canis famill  
P00762 rattus norv  
Q9n2d1 sus scrofa  
P97435 mus musculus  
Q90627 gallus gall  
Q90628 gallus gall  
P98073 homo sapien  
P80009 canis famill  
Q90629 gallus gall  
O15393 homo sapien

ALIGNMENTS

RESULT 1

CTRL_HUMAN	STANDARD;	PRT;	264 AA.
ID	CTRL_HUMAN		
AC	P40313;		
DT	01-FEB-1995 (Rel. 31, Created)		
DT	01-FEB-1995 (Rel. 31, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Chymotrypsin-like protease CTRL-1 precursor (EC 3.4.21.-)		
GN	CTRL OR CTRL1		
OS	Homo sapiens (Human)		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=94093544; PubMed=8268911;		
RA	Larsen F., Solheim J., Kristensen T., Kolsto A.B., Prydz H.;		
RT	"A tight cluster of five unrelated human genes on chromosome		
RT	16q22.1."		
RL	Hum. Mol. Genet. 2:1589-1595(1993).		
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; X71874; CAA50710.1; -		
DR	EMBL; X71877; CAA50711.1; -		
DR	HSSP; P00763; LDPO.		
DR	MEROPS; S01-256; -		
DR	Genew; HGNC:2524; CTRL.		
DR	MIM; 118888; -		
DR	InterPro; IPR001314; Chymotrypsin.		
DR	InterPro; IPR001254; Ser_protease_Try.		
DR	Pfam; PF00089; trypsin; 1.		
DR	PRINTS; PS00722; CHYMOTRYPSIN.		
DR	SMART; SM00020; Tryp_Spc; 1.		
DR	PROSITE; PS02040; TRYPsin_DOM; 1.		
DR	PROSITE; PS00134; TRYPsin_HIS; 1.		
DR	PROSITE; PS00135; TRYPsin_SER; 1.		
KW	Hydrolase; Serine protease; Glycoprotein; zymogen; Signal.		
FT	SIGNAL 1 18		
FT	PROPEP 19 33		
FT	CHAIN 34 264		
FT	ACT_SITE 75 75		
FT	ACT_SITE 121 121		
FT	ACT_SITE 214 214		
FT	CARBOHYD 114 114		
FT	DISULFID 19 141		
FT	DISULFID 60 76		
FT	DISULFID 155 220		
FT	DISULFID 187 201		

POTENTIAL.

ACTIVATION PEPTIDE (POTENTIAL).

CHYMOTRYPSIN-LIKE PROTEASE CTRL-1.

CHARGE RELAY SYSTEM (BY SIMILARITY).

CHARGE RELAY SYSTEM (BY SIMILARITY).

CHARGE RELAY SYSTEM (BY SIMILARITY).

N-LINKED (GLCNAC. . .) (POTENTIAL).

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

FT DISULFID 210 239 BY SIMILARITY.  
SQ SEQUENCE 264 AA; 28002 MW; 3FG29F02FA6DDFB4 CRC64;

Query Match 99.4%; Score 1214; DB 1; Length 264;  
Best Local Similarity 99.6%; Pred. No. 6.8e-98;  
Matches 230; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLLGSGWGGIPAIKPAISQRIYNGENAVLGSWPQVSLQDSSGFHFC 60  
|||||  
DB 1 MLLSLTSLVLLGSGWGGIPAIKPAISQRIYNGENAVLGSWPQVSLQDSSGFHFC 60  
|||||

QY 61 GGLISQSWVTAHCNVSFGRHFVVLGEYDRSSNAEPLQVLSVRAITHPSWNSTMMN 120  
|||||  
DB 61 GGLISQSWVTAHCNVSFGRHFVVLGEYDRSSNAEPLQVLSVRAITHPSWNSTMMN 120  
|||||

QY 121 DVTLLKASPAQYTRISPVCLASSNEALTEGLTCTVTTGWRSLGSGVNTPAHLQOVALP 180  
|||||  
DB 121 DVTLLKASPAQYTRISPVCLASSNEALTEGLTCTVTTGWRSLGSGVNTPAHLQOVALP 180  
|||||

QY 181 LVTVNOCROYWSSITDSMICAGGAGASSCGDGGPLVCQKGTWVLIGI 231  
|||||  
DB 181 LVTVNOCROYWSSITDSMICAGGAGASSCGDGGPLVCQKGTWVLIGI 231  
|||||

RESULT 2  
CTR2\_CANFA  
ID CTR2\_CANFA STANDARD; PRT; 263 AA.  
AC P04813;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Chymotrypsinogen 2 precursor (EC 3.4.21.1).  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=84170253; PubMed=6584866;  
RA Pinsky S.D., Laforge K.S., Luc V., Scheele G.;  
RT Identification of cDNA clones encoding secretory isoenzyme forms:  
RT sequence determination of canine pancreatic chymotrypsinogen 2  
RNA.";  
RL Proc. Natl. Acad. Sci. U.S.A. 80:7486-7490(1983).  
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,  
CC Phe-I-Xaa, Leu-I-Xaa.  
CC -!- SUBCELLULAR LOCATION: Extracellular.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: K01173; AAA30841.1; -  
DR FIR; A21195; A21195.  
DR HSP; P00766; IACB.  
DR MEROPS; S01.152; -  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR001254; Ser.protease\_Try.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; TRYP\_SPC; 1.  
DR PROSITE; PS0240; TRYPsin\_DOM; 1.  
DR PROSITE; PS00134; TRYPsin\_HIS; 1.  
DR PROSITE; PS00135; TRYPsin\_SPR; 1.  
KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.  
FT SIGNAL 1 18  
FT CHAIN 19 263 CHYMOTRYPSINOGEN 2.  
FT CHAIN 19 31 CHYMOTRYPSIN 2, A CHAIN.

FT CHAIN 34 164  
FT CHAIN 167 263  
FT ACT\_SITE 75 75  
FT ACT\_SITE 120 120  
FT ACT\_SITE 213 213  
FT DISULFID 19 140  
FT DISULFID 60 76  
FT DISULFID 154 219  
FT DISULFID 186 200  
FT DISULFID 209 238  
SQ SEQUENCE 263 AA; 27787 MW; 2A2F449D813B3961 CRC64;

Query Match 58.5%; Score 714.5; DB 1; Length 263;  
Best Local Similarity 56.7%; Pred. No. 9.7e-55;  
Matches 131; Conservative 35; Mismatches 64; Indels 1; Gaps 1;

QY 1 MLLSLTSLVLLGSGWGGIPAIKPAISQRIYNGENAVLGSWPQVSLQDSSGFHFC 60  
|||||  
DB 1 MAFLLWLLSCFALLGTAFGCGVPAIQVLSGLSRVINGEDAVPGSWPQVSLQDSTGFHFC 60  
|||||

QY 61 GGLISQSWVTAHCNVSFGRHFVVLGEYDRSSNAEPLQVLSVRAITHPSWNSTMMN 120  
|||||  
DB 61 GGLISQSWVTAHCNVSFGRHFVVLGEYDRSSNAEPLQVLSVRAITHPSWNSTMMN 120  
|||||

QY 121 DVTLLKASPAQYTRISPVCLASSNEALTEGLTCTVTTGWRSLGSGVNTPAHLQOVALP 180  
|||||  
DB 121 DVTLLKASPAQYTRISPVCLASSNEALTEGLTCTVTTGWRSLGSGVNTPAHLQOVALP 180  
|||||

QY 181 LVTVNOCROYWSSITDSMICAGGAGASSCGDGGPLVCQKGTWVLIGI 231  
|||||  
DB 181 LVTVNOCROYWSSITDSMICAGGAGASSCGDGGPLVCQKGTWVLIGI 231  
|||||

RESULT 3  
CTR2\_GADMO  
ID CTR2\_GADMO STANDARD; PRT; 263 AA.  
AC P47796;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Chymotrypsin A precursor (EC 3.4.21.1).  
OS Gadus morhua (Atlantic cod).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.  
OX NCBI\_TaxID=8049;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Pyloric caeca;  
RX MEDLINE=94368860; PubMed=8086467;  
RA Gudmundsdottir A., Oskarsson S., Eakin A.E., Craik C.S.,  
RA Bjarnason J.B.;  
RT "Atlantic cod cDNA encoding a psychrophilic chymotrypsinogen.";  
RL Biochim. Biophys. Acta 1219:211-214(1994).  
RN [2]  
RP SEQUENCE OF 19-30 AND 34-49.  
RX TISSUE=Pyloric caeca;  
RX MEDLINE=92111252; PubMed=1764912;  
RA Asgerlsson B., Bjarnason J.B.;  
RT "Structural and kinetic properties of chymotrypsin from Atlantic cod  
RT (Gadus morhua). Comparison with bovine chymotrypsin.";  
RL Comp. Biochem. Physiol. 99B:327-335(1991).  
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,  
CC Phe-I-Xaa, Leu-I-Xaa.  
CC -!- SUBCELLULAR LOCATION: Extracellular.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

or send an email to license@isb-sib.ch).

```

CC EMBL; X78490; CAA55242.1; -.
CC HSSP; P00766; ICHG.
CC MEROPS; S01.152; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
FT SIGNAL 1 18
FT CHAIN 19 263
FT ACT_SITE 75 75 CHYMOTRYPSIN A.
FT ACT_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 19 140 BY SIMILARITY.
FT DISULFID 60 76 BY SIMILARITY.
FT DISULFID 154 219 BY SIMILARITY.
FT DISULFID 186 200 BY SIMILARITY.
FT DISULFID 209 238 BY SIMILARITY.
FT CONFLICT 21 21 R -> S (IN REF. 2).
FT CONFLICT 25 25 S -> Q (IN REF. 2).
FT CONFLICT 29 29 T -> S (IN REF. 2).
FT CONFLICT 44 44 S -> T (IN REF. 2).
FT CONFLICT 46 46 S -> Y (IN REF. 2).
SQ SEQUENCE 263 AA; 28294 MW; 47AAC699A0A64FB CRC64;

Query Match 57.5%; Score 701.5; DB 1; Length 263;
Best Local Similarity 58.6%; Pred. No. 1.3e-53;
Matches 129; Conservative 33; Mismatches 57; Indels 1; Gaps 1;

Qy 12 LLGSSWGGGIPAIKPAISFQRIYNGENAVLGSMPQVSLQDSSGFHFCGSLISQSWV 71
Db 12 LFRRTYGGCPAISVITGYSRIYNGEAVPHSWSQVSLQDQGTGFHFCGSLINENWV 71
Qy 72 TAAHCNVSPGRHVVGLGYDRSSNAEPLOVLSVRAITHPSWNSNTMNDVTLKLASPA 131
Db 72 TAAHCNVK-NYHRVVLGHDHSSSEGQVMTVGQVFKHPRYNGFTINDILLVKLATPA 130
Qy 132 QYTRISPVCLASNEALTEGLCTVTTGWRSLGSGVNTPAHLQOVALPLTVNQCRQYW 191
Db 131 TLNMRVSPVCLAEETDDVPEGGMKCVTSGMGLTRYNAADTALLQOAAALPLLTNECKRFW 190
Qy 192 DSSITDSMICAGGACSCQDSSGGPLVCQKGNWTLVGI 231
Db 191 GNKISDLMICAGAGACSCMGDSGGPLVCQKAGSWTLVGI 230

```

RESULT 4

```

ID CTRB_HUMAN STANDARD; PRT; 263 AA.
AC P17538;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymotrypsinogen B precursor (EC 3.4.21.1).
GN CTRB1 OR CTRB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=89134264; PubMed=2917002;
RA Tomita N., Izumoto Y., Horii A., Doi S., Yokouchi H., Ogawa M.,
RA Mori T., Matsubara K.;
RT "Molecular cloning and nucleotide sequence of human pancreatic
prechymotrypsinogen cDNA.";

```

```

RL Biochem. Biophys. Res. Commun. 158:569-575(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RA Strausberg R.;
RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,
CC Phe-I-Xaa, Leu-I-Xaa;
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M24400; AAA52128.1; -.
CC EMBL; BC005385; AAH05385.1; -.
CC PIR; A31299; A31299.
CC HSSP; P00766; ICHG.
CC MEROPS; S01.152; -.
CC MIN; I18890; -.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC SMART; SM00020; TRYPSIN_DOM; 1.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
FT SIGNAL 1 18
FT CHAIN 19 263 CHYMOTRYPSINOGEN B.
FT CHAIN 19 31 CHYMOTRYPSIN B, A CHAIN.
FT CHAIN 34 164 CHYMOTRYPSIN B, B CHAIN.
FT CHAIN 167 263 CHYMOTRYPSIN B, C CHAIN.
FT ACT_SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 19 140 BY SIMILARITY.
FT DISULFID 60 76 BY SIMILARITY.
FT DISULFID 154 219 BY SIMILARITY.
FT DISULFID 186 200 BY SIMILARITY.
FT DISULFID 209 238 BY SIMILARITY.
SQ SEQUENCE 263 AA; 27870 MW; 4C1C055A490B8701 CRC64;

```

Query Match 56.5%; Score 689.5; DB 1; Length 263;

Best Local Similarity 55.4%; Pred. No. 1.4e-52;

Matches 128; Conservative 35; Mismatches 67; Indels 1; Gaps 1;

```

Qy 1 MLLSLTSLVLLGSSWGGGIPAIKPAISFQRIYNGENAVLGSMPQVSLQDSSGFHFC 60
Db 1 MAFLWLLSCWALLGTTFGCGVPAIHPLVLSGLSRVNGEDAVPGSPWPQVSLQDQGTGFHFC 60
Qy 61 GGSLSISQSWVYTAHCNVSPGRHVVGLGYDRSSNAEPLOVLSVRAITHPSWNSNTMNN 120
Db 61 GGSLSIEDWVYTAACGVRYS-DVVYVAGEFDQGSDEENIQVLKIAKVFKNPKFSILTNN 119
Qy 121 DVTLLKASPQYTRISPVCLASNEALTEGLCTVTTGWRSLGSGVNTPAHLQOVALP 180
Db 120 DITLLKATPARFSQTVSAVCLPSADDDPPACTLCATGCTGKTKYANKTTPDKLQOALP 179
Qy 181 LVTNQCROYWDSSTDSMICAGGACSCQDSSGGPLVCQKGNWTLVGI 231
Db 180 LLSNAECKKSGRRITDVMICAGAGSCVSSCMGDSGGPLVCQKDGAWTLVGI 230

```

RESULT 5

CTRB\_RAT

ID CTRB\_RAT STANDARD; PRT; 263 AA.  
AC P07338;  
DT 01-APR-1988 (Rel. 07, Created)  
DT 01-APR-1988 (Rel. 07, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Chymotrypsinogen B precursor (EC 3.4.21.1).  
GN CTRB1.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85054881; PubMed=6209274;  
RA Bell G.I., Quinto C., Quiroga M., Valenzuela P., Craik C.S.,  
RA Rutter W.J.;  
RT Isolation and sequence of a rat chymotrypsin B gene.";  
RL J. Biol. Chem. 259:14265-14270(1984).  
CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-|-Xaa, Trp-|-Xaa,  
CC phe-|-Xaa, Leu-|-Xaa.  
CC -|- SUBCELLULAR LOCATION: Extracellular.  
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; K02298; AAA98732.1; -.  
DR PIR; A22658; KYRUB.  
DR HSP; P00766; 1CHG.  
DR MEROPS; S01.152; -.  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR001254; Ser\_protease\_Try.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; TRYP\_SPC; 1.  
DR PROSITE; PS50240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydroxylase; Serine protease; Digestion; Pancreas; Zymogen; Signal.  
FT SIGNAL 1 18  
FT CHAIN 19 263 CHYMOTRYPSINOGEN B.  
FT CHAIN 19 31 CHYMOTRYPSIN B, A CHAIN.  
FT CHAIN 34 164 CHYMOTRYPSIN B, B CHAIN.  
FT CHAIN 167 263 CHYMOTRYPSIN B, C CHAIN.  
FT ACT\_SITE 75 75 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT DISULFID 19 140 BY SIMILARITY.  
FT DISULFID 60 76 BY SIMILARITY.  
FT DISULFID 154 219 BY SIMILARITY.  
FT DISULFID 186 200 BY SIMILARITY.  
FT DISULFID 209 238 BY SIMILARITY.  
SQ SEQUENCE 263 AA; 27849 MW; ACAFDACF8C4DA6D CRC64;  
Query Match 54.9%; Score 670.5; DB 1; Length 263;  
Best Local Similarity 53.2%; Pred. No. 6.1e-51;  
Matches 123; Conservative 35; Mismatches 72; Indels 1; Gaps 1;  
QY 1 MLLSLTLVLGSGWGGCGIPALPALSFSORIVNGENAVLGSWPQVSLQDSSGFHFC 60  
DB 1 MAFLLVSCFALVATFGCGVTIQVLGLSRVNGEDAIPEGSPWPQVSLQDKTGFHFC 60  
QY 61 GGLISQSWVTAHNCVSPGRHVFVLYGDRSSNAEPLQVLSVSRATHPSWNSTMMN 120  
DB 61 GGLISQSWVTAHNCVSPGRHVFVLYGDRSSNAEPLQVLSVSRATHPSWNSTMMN 119  
QY 121 DVTLKLSAPQYTTTRISPVYCLASSNEALTEGLTCVTTCGRLSGVGNVTPAHLQVVALP 180  
DB 121 DVTLKLSAPQYTTTRISPVYCLASSNEALTEGLTCVTTCGRLSGVGNVTPAHLQVVALP 111

DB 120 DITLLKLTAPQAFSETVSAVCLPNVDDPPGCTGTCATGCGTKYNALKTPEKLOQALP 179  
QY 181 LVTVNQCROYWDSITDSMICAGGAGASSCGDSGLVLCQKGNWVLGI 231  
DB 180 IVSEADCKKSGSKITDVTMTAGASSVSSCMGDSGLVLCQKGDVWTLGI 230  
RESULT 6  
ID CTRB\_BOVIN STANDARD; PRT; 245 AA.  
AC P00767;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Chymotrypsinogen B (EC 3.4.21.1).  
OS Bos taurus (Bovine)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE, DISULFIDE BONDS, AND ACTIVE SITE.  
RX MEDLINE=68238908; PubMed=5649671;  
RA Smillie L.B., Furka A., Nagabhushan N., Stevenson K.J., Parkes C.O.;  
RT "Structure of chymotrypsinogen B compared with chymotrypsinogen A and  
RT trypsinogen.";  
RL Nature 218:343-346(1968).  
CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-|-Xaa, Trp-|-Xaa,  
CC phe-|-Xaa, Leu-|-Xaa.  
CC -|- SUBCELLULAR LOCATION: Extracellular.  
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -|- DATABASE: NAME=Washington enzyme manual;  
CC WWW="http://www.worthington-biochem.com/manual/C/CHY.html".  
DR PIR; A00953; KYBOB.  
DR HSP; P00766; 1ACB.  
DR MEROPS; S01.152; -.  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR001254; Ser\_protease\_Try.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; TRYP\_SPC; 1.  
DR PROSITE; PS50240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydroxylase; Serine protease; Digestion; Pancreas; Zymogen.  
FT CHAIN 1 13 CHYMOTRYPSIN B, A CHAIN.  
FT CHAIN 16 146 CHYMOTRYPSIN B, B CHAIN.  
FT CHAIN 149 245 CHYMOTRYPSIN B, C CHAIN.  
FT ACT\_SITE 57 57 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 102 102 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 195 195 CHARGE RELAY SYSTEM.  
FT DISULFID 1 122  
FT DISULFID 42 58  
FT DISULFID 136 201  
FT DISULFID 168 182  
FT DISULFID 191 220  
SQ SEQUENCE 245 AA; 25755 MW; 678016446FF5FEB5 CRC64;  
Query Match 54.6%; Score 666.5; DB 1; Length 245;  
Best Local Similarity 57.3%; Pred. No. 1.3e-50;  
Matches 122; Conservative 28; Mismatches 62; Indels 1; Gaps 1;  
QY 19 CGPIAKPALSPQSRIVNGENAVLGSWPQVSLQDSSGFHFCGGLISQSWVTAHNCV 78  
DB 1 CGVPAIQPVLSGLRVNGEDAVPGSPWPQVSLQDSTGFHFCGGLISQSWVTAHNCV 60  
QY 79 SPGRHFVLYGDRSSNAEPLQVLSVSRATHPSWNSTMMNNDVTLKLSAPQYTTTRIS 138  
DB 61 TTS-DVVVAGEFDQGLEDTQVLKIKVFNKPKFSILTVRNDITLLKLTAPQAFSETVS 119  
QY 139 PYCLASSNEALTEGLTCVTTCGRLSGVGNVTPAHLQVVALPVTNOCROYWDSITDS 198  
DB 120 AVCLPSADEDFPAGMLCATTGKTKYNALKTPEKLOQALPVTNOCROYWDSITDS 179

```
QY 199 MICAGGAGSSCGDGGPLVCOKGNTWVLGI 231
      ||||| : ||| ||||| ||||| |||||
Db 180 MICAGASGVSSCMGDSGGPLVCOKGNTWVLGI 212

RESULT 7
CTRB_GADMO          STANDARD;          PRT;    245 AA.
ID CTRB_GADMO
AC P80646;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymotrypsin B (EC 3.4.21.1).
OS Gadus morhua (Atlantic cod).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
OX NCBI_TaxID=8049;
RN [1]
RP SEQUENCE.
RC TISSUE=Pyloric caeca;
RX MEDLINE=96439045; PubMed=8841380;
RA Leth-Larsen R., Asgeirsson B., Thorolfsson M., Noerregaard-Madsen M.,
RA Hoejrup P.;
RT "Structure of chymotrypsin variant B from Atlantic cod, Gadus
RT morhua.";
RL Biochim. Biophys. Acta 1297:49-56(1996).
RN [2]
RP SEQUENCE OF 1-12 AND 16-31.
RC TISSUE=Pyloric caeca;
RX MEDLINE=92111252; PubMed=1764912;
RA Asgeirsson B., Bjarnason J.B.;
RT "Structural and kinetic properties of chymotrypsin from Atlantic cod
RT (Gadus morhua). Comparison with bovine chymotrypsin.";
RL Comp. Biochem. Physiol. 99B:327-335(1991).
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,
CC Phe-I-Xaa, Leu-I-Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR HSP: P00766; 1CHG.
DR MEROPS: S01.152; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen.
FT CHAIN 1 13 CHYMOTRYPSIN B, A CHAIN.
FT CHAIN 16 245 CHYMOTRYPSIN B, B CHAIN.
FT ACT_SITE 57 57 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 101 101 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 195 195 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 1 121 BY SIMILARITY.
FT DISULFID 42 58 BY SIMILARITY.
FT DISULFID 135 201 BY SIMILARITY.
FT DISULFID 167 182 BY SIMILARITY.
FT DISULFID 191 220 BY SIMILARITY.
FT CONFLICT 9 11 QVT -> VIS (IN REF. 2).
FT CONFLICT 26 26 S -> T (IN REF. 2).
FT CONFLICT 28 29 PW -> Y (IN REF. 2).
SQ SEQUENCE 245 AA; 26260 MW; 74FE0D425517AB02 CRC64;

Query Match 54.4%; Score 664.5; DB 1; Length 245;
Best Local Similarity 58.3%; Pred. No. 1.9e-50;
Matches 126; Conservative 34; Mismatches 49; Indels 7; Gaps 5;

QY 19 CGIPAIPALSFQRIVNGENAVLGSWPQVSLQDSSGFHFCGSLISOSVWVTAACNV 78
      ||||| : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 CGSPAIPQVGTGARIVNGEAVPHSWPQVSLQDSSGFHFCGSLINENWVYVTAACNV 60

QY 79 SPGRFVVLGEYDRSSNAEPLOVLSVRAITHPSWNSTTMNDVTLKLKSPAOVTRTS 138
      . ||||| : ||| : ||| ||||| ||||| ||||| ||||| |||||
Db 61 RT-YHRVIVGEHDKASD-ENIQILKPSVYFTHPKWDSRTINNDISLIKASPAVLGTNVS 118

QY 139 PVLCLASSNEALTEGLTCVTTGNG--RLSGVGNVTPAHLOQVAPLVTVVNOCRQYW-DSIS 195
      ||||| : ||| ||||| : ||| ||||| ||||| ||||| ||||| |||||
Db 119 PVLCESSDVFAFGMKCVTSGMGLTRYNAPG--TPNKLQAAALPLMSNEECSTMGNNMI 176

QY 196 TDSMICAGGAGASSCGDGGPLVCOKGNTWVLGI 231
      : ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 177 SDVMICAGAAGATSCMGDSGGPLVCOKGNTWVLGI 212

RESULT 8
CTRA_BOVIN          STANDARD;          PRT;    245 AA.
ID CTRA_BOVIN
AC P00766;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymotrypsinogen A (EC 3.4.21.1).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE, DISULFIDE BONDS, AND ACTIVE SITE.
RX MEDLINE=67181721; PubMed=5971783;
RA Brown J.R., Hartley B.S.;
RT "Location of disulphide bridges by diagonal paper electrophoresis.
RT The disulphide bridges of bovine chymotrypsinogen A.";
RL Biochem. J. 101:214-228(1966).
RN [2]
RP REVISION TO 102.
RX MEDLINE=69106266; PubMed=5764436;
RA Blow D.M., Birktoft J.J., Hartley B.S.;
RT "Role of a buried acid group in the mechanism of action of
RT chymotrypsin.";
RL Nature 221:337-340(1969).
RN [3]
RP PRELIMINARY SEQUENCE.
RA Hartley B.S.;
RT "Amino-acid sequence of bovine chymotrypsinogen-A.";
RL Nature 201:1284-1287(1964).
RN [4]
RP SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=67183948; PubMed=5972866;
RA Meloun B., Klueh I., Kostka V., Moravsek L., Prusik Z., Vanacek J.,
RA Keil B., Sorm F.;
RT "Covalent structure of bovine chymotrypsinogen A.";
RL Biochim. Biophys. Acta 130:543-546(1966).
RN [5]
RP ACTIVE SITE.
RX MEDLINE=67181723; PubMed=5971785;
RA Smillie L.B., Hartley B.S.;
RT "Histidine sequences in the active centres of some 'serine'
RT proteinases.";
RL Biochem. J. 101:232-241(1966).
RN [6]
RP X-RAY CRYSTALLOGRAPHY.
RX MEDLINE=72035052; PubMed=4399050;
RA Birktoft J.J., Blow D.M., Henderson R., Steitz T.A.;
RT "I. Serine Proteinases. The structure of alpha-chymotrypsin.";
RL Philos. Trans. R. Soc. Lond., B, Biol. Sci. 257:67-76(1970).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF CHYMOTRYPSINOGEN.
RX MEDLINE=7017557; PubMed=5442169;
RA Freer S.T., Kraut J., Robertus J.D., Wright H.T., Xuong N.H.;
RT "Chymotrypsinogen: 2.5-A crystal structure, comparison with alpha-
RT chymotrypsin, and implications for zymogen activation.";
RL Biochemistry 9:1997-2009(1970).
```

[8]	X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF GAMMA-CHYMOTRYPSIN.
RN	MEDLINE=82078042; PubMed=6914398;
RX	Cohen G.H., Silverton E.W., Davies D.R.;
RT	"Refined crystal structure of gamma-chymotrypsin at 1.9-A resolution.
RT	Comparison with other pancreatic serine proteases.";
J.	J. Mol. Biol. 148:449-479(1981).
[9]	
RN	X-RAY CRYSTALLOGRAPHY (1.68 ANGSTROMS) OF ALPHA-CHYMOTRYPSIN.
RX	MEDLINE=86011575; PubMed=4046030;
RT	Tsakada H., Blow D.M.;
RT	"Structure of alpha-chymotrypsin refined at 1.68-A resolution.";
RL	J. Mol. Biol. 184:703-711(1985).
- -	CATALYTIC ACTIVITY: Preferential cleavage: Tyr- -Xaa, Trp- -Xaa,
CC	Phe- -Xaa, Leu- -Xaa.
- -	SUBCELLULAR LOCATION: Extracellular.
CC	- - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
- -	DATABASE: NAMB-Worthington enzyme manual;
CC	WWW=http://www.worthington-biochem.com/manual/C/CHY.html".
PIR	A0952; KBOA.
FDB	2CGA; 15-APR-90.
FDB	2CHA; 31-MAY-84.
FDB	4CHA; 29-OCT-85.
FDB	5CHA; 16-OCT-87.
FDB	6CHA; 16-OCT-87.
FDB	1CHG; 27-JAN-84.
FDB	1CHO; 16-JUL-88.
FDB	2GCH; 31-MAY-84.
FDB	3GCH; 15-OCT-92.
FDB	4GCH; 15-OCT-90.
FDB	5GCH; 15-OCT-90.
FDB	6GCH; 15-OCT-90.
FDB	7GCH; 15-OCT-90.
FDB	8GCH; 15-JUL-93.
FDB	1GCT; 15-OCT-91.
FDB	2GCT; 15-OCT-91.
FDB	3GCT; 15-OCT-91.
FDB	1ACB; 31-OCT-93.
FDB	1GMC; 31-OCT-93.
FDB	1GMD; 31-OCT-93.
FDB	1CGI; 30-APR-94.
FDB	1CGJ; 30-APR-94.
FDB	1GCD; 22-JUN-94.
FDB	1GHA; 22-JUN-94.
FDB	1GHB; 22-JUN-94.
FDB	1GMH; 30-SEP-94.
FDB	2GMT; 01-NOV-94.
FDB	1MTN; 17-AUG-96.
FDB	1AB9; 20-AUG-97.
FDB	1AFQ; 17-SEP-97.
FDB	1CAO; 23-JUL-97.
FDB	1CBW; 23-JUL-97.
FDB	1VGC; 12-NOV-97.
FDB	2VGC; 12-NOV-97.
FDB	3VGC; 12-NOV-97.
FDB	4VGC; 12-NOV-97.
FDB	1HJA; 14-JAN-98.
FDB	MEROPS; S01.001; "
DR	InterPro: IPR001314; Chymotrypsin.
DR	InterPro: IPR001254; Ser.protease_Try.
DR	Pfam: PF00089; trypsin; 1.
DR	PRINTS; PR00722; CHYMOTRYPSIN.
DR	SMART; SM00020; Tryp_Spc; 1.
DR	PROSITE; PS00240; TRYPSIN_DOM; 1.
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.
DR	PROSITE; PS00135; TRYPSIN_SER; 1.
KW	Hydrolase; Serine protease; Digestion; Pancreas; Zymogen;
KW	3D-structure.
FT	CHAIN 1 13 CHYMOTRYPSIN A, A CHAIN.
FT	CHAIN 16 146 CHYMOTRYPSIN A, B CHAIN.
FT	CHAIN 149 245 CHYMOTRYPSIN A, C CHAIN.
FT	ACT_SITE 57 57 CHARGE RELAY SYSTEM.
FT	ACT_SITE 102 102 CHARGE RELAY SYSTEM.

ID	EL2_PIG	STANDARD;	PRT;	269 AA.
AC	P08419;			
DT	01-AUG-1988 (Rel. 08, Created)			
DT	01-AUG-1988 (Rel. 08, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Elastase 2 precursor (EC 3.4.21.71).			
DE	ELA2.			
OS	Sus scrofa (Pig).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
ON	NCBI_TaxID=9823;			
OX	[1]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=87217962; PubMed=3646943;			
RA	Kawashima I., Tani T., Shimoda K., Takiguchi Y.;			
RA	"Characterization of pancreatic elastase II cDNAs: two elastase II			
RT	mRNAs are expressed in human pancreas.";			
RL	NCBI_6163-172(1987).			
RL	[2]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=88198076; PubMed=2834346;			
RA	Shirasu Y., Yoshida H., Matsuki S., Takemura K., Ikeda N.,			
RA	Shimada Y., Ozawa T., Mikayama T., Iijima H., Ishida A., Sato Y.,			
RA	Tamai Y., Tanaka J., Ikenaga H.;			
RT	"Molecular cloning and expression in Escherichia coli of a cDNA			
RT	encoding human pancreatic elastase 2.";			
RL	J. Biochem. 102:1555-1563(1987).			
CC	-1- FUNCTION: ACTS UPON ELASTIN.			
CC	-1- CATALYTIC ACTIVITY: Preferential cleavage: Leu- -Xaa, Met- -Xaa			
CC	and Phe- -Xaa. Hydrolyzes elastin.			
CC	-1- SUBCELLULAR LOCATION: Secreted.			
CC	-1- TISSUE SPECIFICITY: PANCREAS.			
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@lsb-sib.ch">license@lsb-sib.ch</a> ).			
CC	-----			
DR	EMBL; M16651; AAA31027.1; -;			
DR	EMBL; D00237; BAA00166.1; -;			
DR	PIR; A26823; A26823.			
DR	HSP; P00763; LDPO.			
DR	MEROPS; S01.155; -;			
DR	InterPro; IPR001314; Chymotrypsin.			
DR	InterPro; IPR001254; Ser_protease_Try.			
DR	Pfam; PF00089; trypsin; 1.			
DR	PRINTS; PR00722; CHYMOTRYPSIN.			
DR	SMART; SM0020; TRYP_SPC; 1.			
DR	PROSITE; PS0240; TRYPSIN_DOM; 1.			
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.			
DR	PROSITE; PS00135; TRYPSIN_SER; 1.			
KW	Hydrolase; Serine protease; zymogen; Signal.			
FT	SIGNAL	1	16	
FT	PROPEP	17	28	ACTIVATION PEPTIDE.
FT	CHAIN	29	269	ELASTASE 2.
FT	ACT_SITE	73	73	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE	121	121	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE	216	216	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	DISULFID	58	74	BY SIMILARITY.
FT	DISULFID	155	222	BY SIMILARITY.
FT	DISULFID	186	202	BY SIMILARITY.
FT	DISULFID	212	243	BY SIMILARITY.
FT	CONFLICT	10	10	L -> S (IN REF. 2).
FT	CONFLICT	118	118	N -> K (IN REF. 2).
FT	CONFLICT	132	132	S -> Y (IN REF. 2).
FT	CONFLICT	172	172	I -> V (IN REF. 2).
FT	CONFLICT	202	202	C -> V (IN REF. 2).
FT	SEQUENCE	269 AA;	28699 MW;	BACGFG9AF4DDE56 CRC64;

```

CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Ala-Xaa. Does not
CC hydrolyse elastin.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
CC -!- CAUTION: Was originally (Ref.5) thought to be elastase 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M16630; AAA36482.1; -.
CC EMBL: BC005216; AAH05216.1; -.
CC EMBL: M18692; AAA58454.1; -.
CC PIR: B29934; B29934.
CC PIR: A27206; A27206.
CC PIR: S04999; S04999.
CC PIR: S04490; S04490.
CC HSP: P05805; IFON.
CC MEROPS: S01.205; -.
CC GlycoSuiteDB; P08861; -.
CC SWISS-2DPAGE; P08861; HUMAN.
CC Genew; HGNC:15945; ELA3B.
CC InterPro; IPR001314; Chymotrypsin.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC SMART; SM00020; TRYP_SPC; 1.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC Hydrolase; Serine protease; Zymogen; Signal; Glycoprotein.
CC OR 16 (POTENTIAL).
CC SIGNAL 1 15
CC PROPEP 16 28
CC CHAIN 29 270
CC ELASTASE IIIB.
CC ACT. SITE 73 73
CC CHARGE RELAY SYSTEM (BY SIMILARITY).
CC ACT. SITE 123 123
CC CHARGE RELAY SYSTEM (BY SIMILARITY).
CC ACT. SITE 217 217
CC CHARGE RELAY SYSTEM (BY SIMILARITY).
CC DISULFID 58 74
CC PROBABLE.
CC DISULFID 117 120
CC BY SIMILARITY.
CC DISULFID 157 223
CC BY SIMILARITY.
CC DISULFID 188 204
CC BY SIMILARITY.
CC DISULFID 213 244
CC BY SIMILARITY.
CC CARBOHYD 114 114
CC N-LINKED (GLCNAC. . .).
CC /FTIG-CAR_000212.
CC R -> G (IN REF. 3).
CC A -> G (IN REF. 3).
CC W -> R (IN REF. 1).
CC MISSING (IN REF. 5).
CC R -> P (IN REF. 3).
CC B14BE0AD3695AFE CRC64;
CC SEQUENCE 270 AA; 29293 MW; 29293 MW; B14BE0AD3695AFE CRC64;
Query Match 37.1%; Score 453.5; DB 1; Length 270;
Best Local Similarity 42.5%; Pred. No. 3.5e-32;
Matches 102; Conservative 36; Mismatches 87; Indels 15; Gaps 8;
QY 2 LLLSLTSLVLLGSSGCGIPAIKPAKSPQRIVNGENAVLGSWPQVLSQ-DSSG--FH 58
DB 1 MMLRLSLLLVAVASGYPGPPSSRP-----SSRVNGEDAVPYSPWPQVLSQYKSGSFYH 56
QY 59 FCGSLISQSWVYTAHCNVSPGRHFVVLGEYDRSSNAEPLOVLSVSR--ITHPSWNST 116
DB 57 TCGSLIADPWVYTAGHICSSWTQVVLGEYDRAVKEGPEQVPIPIVNSGDLFVHPLWNKS 116
QY 117 TM--NNDVTLLKLSAQAQYTRISPSVCLASSNEALTEGLTCVTGTGWRLSGGVGNVTPAHL 174
DB 117 CVACGNDIALIKLSRAQLGDVAQLASLPAGDILFNETPCYITGWRLYTNGPL-PDKL 175
QY 175 QQVALPLVTNQCQR--YDSSITDSMICAGGAGACSCGDSGGPLVC-QKGNWTWVLIGI 231
DB 176 QEALLPVVDYEHCSRWNWSSVYKTMVCAGGDIRSGCNGDSGGLPNCPTDGGWQVHG 235

```

```

-RESULT 11
EL2_BOVIN STANDARD: PRT; 269 AA.
ID EL2_BOVIN
AC Q29461;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Elastase 2 precursor (EC 3.4.21.71).
GN ELA2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=98079203; PubMed=9418008;
RA Gestin M., le Huerou-Luron I., Wicker-Planquart C., le Drean G.,
RA Chaix J.C., Puigserver A., Guilloteau P.;
RA "Bovine pancreatic preproelastases I and II: comparison of nucleotide
RA and amino acid sequences and tissue specific expression.";
RL Comp. Biochem. Physiol. 118B:181-187(1997).
CC -!- FUNCTION: ACTS UPON ELASTIN.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Leu-Xaa, Met-Xaa
CC and Phe-Xaa. Hydrolyzes elastin.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: PANCREAS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X97635; CAA66231.1; -.
CC HSP: P00766; ICHG.
CC MEROPS: S01.155; -.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC SMART; SM00020; TRYP_SPC; 1.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC Hydrolase; Serine protease; Zymogen; Signal.
CC SIGNAL 1 16
CC BY SIMILARITY.
CC PROPEP 17 28
CC ACTIVATION PEPTIDE (BY SIMILARITY).
CC CHAIN 29 269
CC ELASTASE 2.
CC ACT. SITE 73 73
CC CHARGE RELAY SYSTEM (BY SIMILARITY).
CC ACT. SITE 121 121
CC CHARGE RELAY SYSTEM (BY SIMILARITY).
CC ACT. SITE 216 216
CC CHARGE RELAY SYSTEM (BY SIMILARITY).
CC DISULFID 58 74
CC BY SIMILARITY.
CC DISULFID 155 222
CC BY SIMILARITY.
CC DISULFID 186 202
CC BY SIMILARITY.
CC DISULFID 212 243
CC BY SIMILARITY.
CC SEQUENCE 269 AA; 28856 MW; 8343B97062CF267C CRC64;
Query Match 37.1%; Score 453; DB 1; Length 269;
Best Local Similarity 43.9%; Pred. No. 3.9e-32;
Matches 105; Conservative 30; Mismatches 90; Indels 14; Gaps 8;
QY 2 LLLSLTSLVLLGSSGCGIPAIKPAKSPQRIVNGENAVLGSWPQVLSQDSSG---FH 58
DB 1 MTRALLSLTVAG-ALSCGVPPYPPOLS---RVVGEDARPNSWPQVLSQYSSSQWRH 56
QY 59 FCGSLISQSWVYTAHCNVSPGRHFVVLGEYDRSSNAEPLOVLSVSR--ITHPSWNST 118

```



```
Db 57 TCGSLIEQNVLTAACHICSSRTYRVVGRQSLSTVESGSLTIAVSKSVIHEKWNSQL 116
Qy 119 --NDVTLKASPAQYTRISPCVCLASSNEALTEGLTCVTTWGRLSGVGNVTPAHLOQ 176
Db 117 AGNDIALKKLASSVPLFDKIQGLCPAAGTILPNNYCYVTWGRQLQSG-ALPDILQ 175
Qy 177 VALPLVTYNQCR--QYDSSITDSMICAGGAG-ASSCGDGGPLVLCOKGN-TWVLIGI 231
Db 176 GKLLVDYATCSNFWGSGTKTNMICAGGDVTSNCGDGGPLNCAANRQWQVHGI 234

RESULT 12
EL2A_HUMAN STANDARD; PRT; 269 AA.
AC P08217; Q14243;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Elastase 2A precursor (EC 3.4.21.71).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87217962; PubMed=3646943;
RA Kawashima I., Tani T., Shimoda K., Takiguchi Y.;
RT "Characterization of pancreatic elastase II cDNAs: two elastase II
RT mRNAs are expressed in human pancreas.";
RL DNA 6:163-172(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88107669; PubMed=3427074;
RA Fletcher T.S., Shen W.F., Largman C.;
RT "Primary structure of human pancreatic elastase 2 determined by
RT sequence analysis of the cloned mRNA.";
RL Biochemistry 26:7256-7261(1987).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Pancreas;
RX MEDLINE=88198076; PubMed=2834346;
RA Shirasu Y., Yoshida H., Matsuki S., Takemura K., Ikeda N., Shimada Y.,
RA Ozawa T., Mikayama T., Iijima H., Ishida A., Sato Y., Tamai Y.,
RA Tanaka J., Ikenaga H.;
RT "Molecular cloning and expression in Escherichia coli of a cDNA
RT encoding human pancreatic elastase 2.";
RL J. Biochem. 102:1555-1563(1987).
RN [4]
RP SEQUENCE FROM N.A.
RA Thomas D.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RX TISSUE=Pancreas;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ACTS UPON ELASTIN.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Leu-I-Xaa, Met-I-Xaa
CC and Phe-I-Xaa. Hydrolyzes elastin.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; M16631; AAA52374.1; -.
```

```
DR EMBL; M16652; AAA52380.1; -
DR EMBL; D00236; BAA00165.1; -
DR EMBL; AL512883; CAC42421.1; -
DR EMBL; BC007031; AA07031.1; -
DR PIR; A27432; A27432.
DR PIR; B26823; B26823.
DR HSSP; P00772; 1ELG.
DR MEROPS; S01.155; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hydrolase; Serine protease; zymogen; Signal.
FT SIGNAL 1 16
FT PROPEP 17 28
FT CHAIN 29 269
FT DISULFID 58 74
FT DISULFID 155 222
FT DISULFID 186 202
FT DISULFID 212 243
FT ACT_SITE 73 73
FT ACT_SITE 121 121
FT ACT_SITE 216 216
FT ACT_SITE 202 202
FT CONFLICT 202 202
SQ SEQUENCE 269 AA; 28888 MW; A2E05143EFF4987C CRC64;

Query Match 36.4%; Score 445; DB 1; Length 269;
Best Local Similarity 43.9%; Pred. No. 1.9e-31;
Matches 105; Conservative 28; Mismatches 92; Indels 14; Gaps 8;

Qy 2 LLSLTSLVLLGSGWCGIPAIKPAISQRIVANGENAVLGSWPQVSLQDSSG---FH 58
Db 1 MIRTLLTLVAG-ALSCGDPTYP--VYTRVVGEEARPNPWPQVSLQYSSNGKQWYH 56
Qy 59 FCGSLISQSWVYTAACHNVSFGRHFVVLGEYDRSSNAEPLQVLSVRAITHPSNSTM 118
Db 57 TCGSLIANSVLTAACHICSSRTYRVGLGRHNLVVAESGLAVSVSVIVHDKWNSQI 116
Qy 119 N--NDVTLKASPAQYTRISPCVCLASSNEALTEGLTCVTTWGRLSGVGNVTPAHLOQ 176
Db 117 SKGNDIALKLANPVSLTDKIQGLCPAAGTILPNNYCYVTWGRQLQNGAV-PDVLOQ 175
Qy 177 VALPLVTYNQCRQ--YDSSITDSMICAGGAG-SSCGDGGPLVLCOKGN-TWVLIGI 231
Db 176 GKLLVDYATCSNFWGSGTKTNMICAGGDVTSNCGDGGPLNCAANRQWQVHGI 234

RESULT 13
EL2_RAT
ID EL2_RAT STANDARD; PRT; 271 AA.
AC P00774;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Elastase 2 precursor (EC 3.4.21.71).
GN ELA2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82182967; PubMed=6918221;
RA McDonald R.J., Swift G.H., Quinto C., Swain W., Pictet R.L.,
RA Nikovits W., Rutter W.J.;
RT "Primary structure of two distinct rat pancreatic preproelastases
RT determined by sequence analysis of the complete cloned messenger
RT ribonucleic acid sequences.";
RL Biochemistry 21:1453-1463(1982).
```

[illegible]

Qy 119 N--NDVTLKLLASPAQYTRISPVCLASSNEALTEGLTCVTGMRGLSGVGNVTPAHQQ 176  
Db 117 SKGNDIALKLANVSLTDKIQACLPAGTILPNPNPCYVTGMRLOTNG-ALPDDLUKQ 175  
Qy 177 VALPLVTVNOCRQ--YWDSSITDSMICAGGAGA--SSCGDSGGPLVCQKGN-TWVLIGI 231  
Db 176 GRLLVVDYATCSSGOWGSTVKTNNICAGGDSVICTNGDSGGPLNCQASDGRWEVHGI 234

## RESULT 15

EL3A\_HUMAN EL3A\_HUMAN STANDARD; PRT; 270 AA.  
AC P09093; O9BRW4;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Elastase IIIA precursor (EC 3.4.21.70) (Protease E).  
GN ELA3.

OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Pancreas;  
RX MEDLINE=88087253; PubMed=2826474;  
RA Tanl T., Ohsumi J., Mita K., Takiguchi Y.;

RT "Identification of a novel class of elastase isozyme, human  
pancreatic elastase III, by cDNA and genomic gene cloning.";  
RL J. Biol. Chem. 263:1231-1239(1988).  
RN [2]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Pancreas;  
RX MEDLINE=89034017; PubMed=2460440;  
RA Shirasu Y., Takemura K., Yoshida H., Sato Y., Iijima H.,  
RA Shimada Y., Miyayama T., Ozawa T., Ikeda N., Ishida A., Tamai Y.,  
RA Matsuki S., Tanaka J., Ikenaga H., Ogawa M.;

RT "Molecular cloning of complementary DNA encoding one of the human  
pancreatic protease E isozymes.";  
RL J. Biochem. 104:259-264(1988).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Prostate;  
RA Strausberg R.;

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: EFFICIENT PROTEASE WITH ALANINE SPECIFICITY BUT ONLY  
CC LITTLE ELASTOLYTIC ACTIVITY.  
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Ala-I-Xaa. Does not  
CC hydrolyse elastin.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; M18700; AAA66350.1; -  
DR EMBL; M18693; AAA66350.1; JOINED.  
DR EMBL; M18694; AAA66350.1; JOINED.  
DR EMBL; M18695; AAA66350.1; JOINED.  
DR EMBL; M18696; AAA66350.1; JOINED.  
DR EMBL; M18697; AAA66350.1; JOINED.  
DR EMBL; M18698; AAA66350.1; JOINED.  
DR EMBL; M18699; AAA66350.1; JOINED.  
DR EMBL; D00306; BAA00212.1; -  
DR EMBL; BC005918; AAH05918.1; -  
DR PIR; A29934; A29934.  
DR HSP; P03805; IFON.  
DR MEROPS; S01.154; -.

DR Genew: HGNC:15944; ELA3A.  
DR InterPro: IPR001314; Chymotrypsin.  
DR InterPro: IPR001254; Ser\_protease\_Try.  
DR Pfam: PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; Tryp\_Spc; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolyase; Serine protease; zymogen; Signal; Glycoprotein.  
FT SIGNAL 1 15  
FT CHAIN 16 28  
FT CHAIN 19 270  
FT ACT\_SITE 73 73  
FT ACT\_SITE 123 123  
FT ACT\_SITE 217 217  
FT DISULFID 58 74  
FT DISULFID 117 120  
FT DISULFID 157 223  
FT DISULFID 188 204  
FT DISULFID 213 244  
FT CARBOHYD 114 114  
FT CONFLICT 63 63  
FT CONFLICT 106 106  
FT CONFLICT 174 174  
SQ SEQUENCE 270 AA; 29474 MW; 576DDB255A4A118C CRC64;

Query Match 34.4%; Score 420.5; DB 1; Length 270;

Best Local Similarity 40.7%; Pred. No. 2.5e-29;

Matches 98; Conservative 35; Mismatches 91; Indels 17; Gaps 8;

Qy 1 MLLSLTLTLVLGLSGGCGGIPATKPAKLSFSORTIVNGENAVLGSPWQVSLQ-DSSG--F 57  
Db 2 MLRLSSLLLVAVASGYG-----PPSHSSRVVHGEDAVPYSPWQVSLQYKSGSFY 55  
Qy 58 HFCGGSLLISQSWVYTAACHNVSPGRHFVYVLGEYDRSSNAEPLQVLSVS--RAITHPSWNS 115  
Db 56 HTC GGSLLIAPDWVYTAGHCISRDLTYQVVLGEYNLAVKEGPEQVTPINSEELFVHPLNWR 115  
Qy 116 TTM--NDVTLKLLASPAQYTRISPVCLASSNEALTEGLTCVTGMRGLSGVGNVTPAH 173  
Db 116 SCVACGNDIALIKLSRSALQGLDVALPAGDILPNKTCYITGWRGLYTNGLP-PDK 174  
Qy 174 LQVVALPLVTVNOCRQ--YWDSSITDSMICAGGAGSCQDSGGPLVC-QKGN TWVLIG 230  
Db 175 LQVRLPVDYKHCSSRWNNWGSTVKKTWVCAGGYIRSCNCGDSGGPLNCPTDGGQVYHG 234  
Qy 231 I 231  
Db 235 V 235

Search completed: December 20, 2002, 15:17:12

Job time : 8.91304 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:16:49 ; Search time 1.04348 seconds  
(without alignments)  
530.147 Million cell updates/sec

Title: US-09-856-319B-4\_COPY\_1\_33

Perfect score: 166

Sequence: 1 MLLSLTSLVLLGSSWCGVPATPALSYNQR 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 106657 seqs, 16763532 residues

Total number of hits satisfying chosen parameters: 106657

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US05\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	133	80.1	192	10	US-09-925-297-529
2	68	41.0	263	10	US-09-888-615-96
3	67	40.4	146	10	US-09-925-297-643
4	55	33.1	231	9	US-10-095-449-2
5	55	33.1	231	10	US-09-448-378-2
6	55	33.1	231	10	US-09-983-806-2
7	55	33.1	242	10	US-09-815-242-13730
8	50	30.1	344	10	US-09-992-647-12
9	48	28.9	13	12	US-10-036-371-5
10	48	28.9	331	9	US-09-860-670-118
11	48	28.9	331	10	US-09-764-853-671
12	48	28.9	331	10	US-09-764-898-219
13	48	28.9	652	10	US-09-992-647-1
14	48	28.9	661	10	US-09-764-853-679
15	48	28.9	661	10	US-09-764-898-224
16	47.5	28.6	20	9	US-10-010-928-7
17	47.5	28.6	20	10	US-09-802-124-7
18	47.5	28.6	20	10	US-09-733-605-7
19	47.5	28.6	78	9	US-09-992-598-363

20	47.5	28.6	78	9	US-09-989-293A-363	Sequence 363, App
21	47.5	28.6	78	9	US-10-063-547-66	Sequence 66, Appl
22	47.5	28.6	78	9	US-09-092-296-15	Sequence 15, Appl
23	47.5	28.6	78	10	US-09-989-722-363	Sequence 363, App
24	47.5	28.6	78	10	US-09-989-723-363	Sequence 363, App
25	47.5	28.6	78	10	US-09-989-279-363	Sequence 363, App
26	47.5	28.6	78	10	US-09-989-727-363	Sequence 363, App
27	47.5	28.6	78	10	US-09-989-731-363	Sequence 363, App
28	47.5	28.6	78	10	US-09-989-732-363	Sequence 363, App
29	47.5	28.6	78	10	US-09-991-073-363	Sequence 363, App
30	47.5	28.6	78	10	US-09-990-442-363	Sequence 363, App
31	47.5	28.6	78	10	US-09-991-163-363	Sequence 363, App
32	47.5	28.6	78	10	US-09-993-604-363	Sequence 363, App
33	47.5	28.6	78	10	US-09-990-456-363	Sequence 363, App
34	47.5	28.6	78	10	US-09-989-721-363	Sequence 363, App
35	47.5	28.6	78	12	US-10-006-867-66	Sequence 66, Appl
36	47.5	28.6	78	12	US-10-052-586-242	Sequence 242, App
37	47.5	28.6	180	9	US-10-010-928-2	Sequence 2, Appl
38	47.5	28.6	180	10	US-09-802-124-2	Sequence 2, Appl
39	47.5	28.6	180	10	US-09-733-605-2	Sequence 2, Appl
40	47	28.3	237	10	US-09-746-284-1	Sequence 1, Appl
41	47	28.3	413	10	US-09-815-242-13998	Sequence 13998, A
42	47	28.3	445	9	US-09-992-598-177	Sequence 177, App
43	47	28.3	445	9	US-09-989-293A-177	Sequence 177, App
44	47	28.3	445	9	US-10-063-547-32	Sequence 32, Appl
45	47	28.3	445	10	US-09-989-722-177	Sequence 177, App

# ALIGNMENTS

RESULT 1  
US-09-925-297-529  
; Sequence 529, Application US/09925297  
; Patent No. US20020081659A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA105  
; CURRENT APPLICATION NUMBER: US/09/925,297  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05989  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 928  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 529  
; LENGTH: 192  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-925-297-529

Query Match 80.1%; Score 133; DB 10; Length 192;  
Best Local Similarity 86.2%; Pred. No. 2e-10;  
Matches 25; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 5 SUTLSLVLLGSSWCGVPATPALSYNQR 33  
|||||  
Db 1 SUTLSLVLLGSSWCGGIPAKPALSFSQR 29  
|||||

RESULT 2  
US-09-888-615-96  
; Sequence 96, Application US/09888615  
; Patent No. US20020064856A1  
; GENERAL INFORMATION:  
; APPLICANT: PLOWMAN, GREGORY  
; APPLICANT: WHYTE, DAVID  
; APPLICANT: CAENEPEL, SEAN  
; APPLICANT: CHARYDCZAK, GLEN  
; APPLICANT: MANNING, GERARD  
; APPLICANT: SUDARSANAM, SUCHA

```

; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038600/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 96
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-888-615-96

```

Query Match 41.0%; Score 68; DB 10; Length 263;  
Best Local Similarity 51.7%; Pred. No. 0.073;  
Matches 15; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

**QY** 1 MLLLSLTSLVLLGSSWCGVPATPALS 29  
| | | : : : : :  
**D6** 1 MASLWLLSCFSLVGAAFGCVPAIHPLVS 29

### RESULT 3

```

US/09-525-2597-043
/ Sequence 643, Application US/099252597
/ Patent No. US20020081659A1
/ GENERAL INFORMATION:
/
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, P
/ FILE REFERENCE: PA105
/ CURRENT APPLICATION NUMBER: US/09/92
/ CURRENT FILING DATE: 2001-08-10
/ PRIOR APPLICATION NUMBER: PCT/US00/0
/ PRIOR FILING DATE: 2000-03-08
/ PRIOR APPLICATION NUMBER: 60/124,270
/ PRIOR FILING DATE: 1999-03-12
/ NUMBER OF SEQ ID NOS: 928
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 643
/ LENGTH: 146

```

```

/ ORGANISM: Homo sapiens
/
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (2)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (34)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (126)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (130)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (133)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (137)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (143)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ US-09-925-297-643

```

Query Match 40.4%; Score 67; DB 10; Length 146;  
Best Local Similarity 66.7%; Pred. No. 0.053;  
Matches 12; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

**Qy**      12 LLGSSWCGVPALTALS 29  
          | : | : | : | : | : | :

Db 16 LVGAAEGCGVPAIHPVLS 33

## RESULT 4

```

US-10-095-449-2
: Sequence 2, Application US/10095449
: Patent No. US20020160004A1
: GENERAL INFORMATION:
: APPLICANT: Lyman, Stewart D.
: Beckmann, M. Patricia
: TITLE OF INVENTION: Ligands for flt3/flk-2 Receptors
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Stephen L. Malaska, Immunex Corporation
: STREET: 51 University Street
: CITY: Seattle
: STATE: Washington
: COUNTRY: US
: ZIP: 98101
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: Apple Macintosh
: OPERATING SYSTEM: Macintosh 7.0.1
: SOFTWARE: Microsoft Word, Version #5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/10/095,449
: FILING DATE: 13-Mar-2002
: CLASSIFICATION DATA:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/669,692
: FILING DATE: 24-JUN-1996
: APPLICATION NUMBER: US/08/162,407
: FILING DATE: December 3, 1993
: APPLICATION NUMBER: 08/111,758
: FILING DATE: August 25, 1993
: APPLICATION NUMBER: 08/106,463
: FILING DATE: August 12, 1993
: APPLICATION NUMBER: 08/068,394
: FILING DATE: May 24, 1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Malaska, Stephen L.
: REGISTRATION NUMBER: 32,655
: REFERENCE/DOCKET NUMBER: 2813-C
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 587-0430
: TELEFAX: (206) 233-0644
: TELEX: 756822

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match 33.1%; Score 55; DB 9; Length 231;  
Best Local Similarity 61.1%; Pred. No. 3.1;  
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MLLSLTSLVLLGSSWG 18  
 :||| | | :||| :||  
Db 192 LLLLLPLTLVLLAAAWG 209

## RESULT 5

US-09-448-378-2 ; Sequence 2, Application US/09448378  
; Patent No. US20020034517A1  
; GENERAL INFORMATION:  
; APPLICANT: Brasel, Kenneth  
; TITLE OF INVENTION: Dendritic Cell Stimulatory Factor  
; FILE REFERENCE: 2836-D

;; CURRENT APPLICATION NUMBER: US/09/448,378  
;; \* CURRENT FILING DATE: 1999-11-23  
;; \* NUMBER OF SEQ ID NOS: 2  
;; SOFTWARE: PatentIn version 3.0  
;; SEQ ID NO 2  
;; LENGTH: 231  
;; TYPE: PRF  
;; ORGANISM: Mus sp.  
US-09-448-378-2

Query Match 33.1%; Score 55; DB 10; Length 231;  
Best Local Similarity 61.1%; Pred. No. 3.1;  
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Oy 1 MLLSLTSLVLLGSSWG 18  
: ||| | : ||| : ||  
Db 192 LLLLLPLTLVLLAAWG 209

## RESULT 6

US-09-983-806-2  
; Sequence 2, Application US/09983806  
; Patent NO. US20020107365A1  
; GENERAL INFORMATION:  
; APPLICANT: Lyman, Stewart D.  
; Beckmann, M. Patricia  
; TITLE OF INVENTION: Ligands for flt3/flk-2 Receptors  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Stephen L. Malaska, Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98101

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh 7.0.1  
SOFTWARE: Microsoft Word, Version #5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/983,806  
FILING DATE: 25-Oct-2001  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/444,626  
FILING DATE: 19-MAY-1995

APPLICATION NUMBER: US/08/162,407  
FILING DATE: 03-DEC-1993  
APPLICATION NUMBER: 08/111,758  
FILING DATE: August 25, 1993  
APPLICATION NUMBER: 08/106,463  
FILING DATE: August 12, 1993  
APPLICATION NUMBER: 08/068,394  
FILING DATE: May 24, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Malaska, Stephen L.  
REGISTRATION NUMBER: 32,655  
REFERENCE/DOCKET NUMBER: 2813-C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 587-0430  
TELEFAX: (206) 233-0644  
TELEX: 756822

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 231 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-983-806-2  
Query Match 33.1%; Score 55; DB 10; Length 231;

Best Local Similarity 61.1%; Pred. No. 3.1;  
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Oy 1 MLLSLTSLVLLGSSWG 18  
: ||| | : ||| : ||  
Db 192 LLLLLPLTLVLLAAWG 209

## RESULT 7

US-09-815-242-13730  
; Sequence 13730, Application US/09815242  
; Patent NO. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13730  
; LENGTH: 242  
; TYPE: PRF  
; ORGANISM: Salmonella typhi  
US-09-815-242-13730

Query Match 33.1%; Score 55; DB 10; Length 242;  
Best Local Similarity 58.8%; Pred. No. 3.3;  
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Oy 17 WCGGVPATPALSYNOR 33  
: ||| | : ||| : ||  
Db 92 WGFDPVPAIEKALAHNER 108

## RESULT 8

US-09-992-647-12  
; Sequence 12, Application US/09992647  
; Patent NO. US20020146767A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Hong  
; APPLICANT: Cohan, Victoria L.  
; APPLICANT: Stuart, Susan G.  
; TITLE OF INVENTION: HUMAN EMRL1-LIKE G PROTEIN-COUPLED RECEPTOR  
; FILE REFERENCE: PC-0052 CIP  
; CURRENT APPLICATION NUMBER: US/09/992,647  
; CURRENT FILING DATE: 2002-11-13  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PERL Program  
; SEQ ID NO 12  
; LENGTH: 344

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Genbank ID NO. US20020146767A1 g2935597
US-09-992-647-12

Query Match          30.1%; Score 50; DB 10; Length 344;
Best Local Similarity 37.0%; Pred. No. 21;
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLGGSWGCGVPAITPA 27
   || | : || : | : ||
Db 250 MIAFKATAQLFILGCTWCLGILQVGPA 276

RESULT 9
US-10-036-371-5
; Sequence 5, Application US/10036371
; Patent No. US20020141987A1
; GENERAL INFORMATION:
; APPLICANT: BJARNARSON, JON B.
; TITLE OF INVENTION: FISH SERINE PROTEINASES AND THEIR PHARMACEUTICAL AND
; TITLE OF INVENTION: COSMETIC USE
; FILE REFERENCE: 81691/284960
; CURRENT APPLICATION NUMBER: US/10/036.371
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 09/411,688
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 5086/99
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Bovine sp.
US-10-036-371-5

Query Match          28.9%; Score 48; DB 12; Length 13;
Best Local Similarity 81.8%; Pred. No. 1,2;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 19 CGVPAITPALS 29
   ||||| |||
Db 1 CGVPAIQPVLS 11

RESULT 10
US-09-860-670-118
; Sequence 118, Application US/09860670
; Patent No. US20020165137A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA127P1
; CURRENT APPLICATION NUMBER: US/09/860.670
; CURRENT FILING DATE: 2001-05-21
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 118
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-860-670-118

Query Match          28.9%; Score 48; DB 9; Length 331;
Best Local Similarity 37.0%; Pred. No. 37;
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLGGSWGCGVPAITPA 27
   || | : || : | : ||
Db 228 MIAFKATAQLFILGCTWCLGILQVGPA 254

```



Search completed: December 20, 2002, 15:21:07  
Job time : 1.04348 secs

```
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID NO. US20020146767A1 429905
US-09-992-647-1

Query Match      28.9%; Score 48; DB 10; Length 652;
Best Local Similarity 37.0%; Pred. No. 75;
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLLGSSWCGVPAITPA 27
   || | | | | | | | | | |
Db 549 MIAFRATAQLFILGCTWCLGLLQVGPA 575

RESULT 14
US-09-764-853-679
; Sequence 679, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJZ06
; CURRENT APPLICATION NUMBER: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 679
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-853-679

Query Match      28.9%; Score 48; DB 10; Length 661;
Best Local Similarity 37.0%; Pred. No. 76;
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLLGSSWCGVPAITPA 27
   || | | | | | | | | | |
Db 558 MIAFRATAQLFILGCTWCLGLLQVGPA 584

RESULT 15
US-09-764-898-224
; Sequence 224, Application US/09764898
; Patent No. US20020090673A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJZ01
; CURRENT APPLICATION NUMBER: US/09/764,898
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 311
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 224
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-898-224

Query Match      28.9%; Score 48; DB 10; Length 661;
Best Local Similarity 37.0%; Pred. No. 76;
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLLGSSWCGVPAITPA 27
   || | | | | | | | | | |
Db 558 MIAFRATAQLFILGCTWCLGLLQVGPA 584
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:14:53 ; Search time 1.86957 Seconds  
(without alignments)  
519.349 Million cell updates/sec

Title: US-09-856-319B-4\_COPY\_1\_33

Perfect score: 166

Sequence: 1 MLLSLTSLVLLGSSWCGVPAITPALSYNQR 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA.\*

1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep.\*

2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep.\*

3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep.\*

4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep.\*

5: /cgn2\_6/ptodata/1/1aa/PCRUS\_COMB.pep.\*

6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55.5	33.4	441	3	US-09-052-778-14
2	55	33.1	231	1	US-08-220-379B-7
3	55	33.1	231	1	US-08-243-545-2
4	55	33.1	231	2	US-08-993-962-2
5	55	33.1	231	4	US-09-160-841-2
6	55	33.1	231	5	PCT-US94-05365-2
7	55	33.1	231	5	PCT-US95-03866-6
8	55	33.1	268	1	US-08-270-584A-2
9	55	33.1	268	3	US-08-765-192-2
10	55	33.1	268	3	US-09-199-793-2
11	52	31.3	306	4	US-09-386-642-53
12	52	31.3	319	4	US-09-386-642-12
13	50	30.1	344	3	US-09-110-116-4
14	48.5	29.2	280	4	US-09-296-284-27
15	48.5	29.2	306	4	US-09-296-284-6
16	48	28.9	521	4	US-08-956-322-4
17	48	28.9	652	3	US-09-110-116-1
18	48	28.9	652	4	US-08-956-322-2
19	47	28.3	838	2	US-08-696-944-19
20	47	28.3	1620	1	US-08-542-363-2
21	47	28.3	1620	4	US-09-100-089-2
22	47	28.3	1620	4	US-09-670-827-2
23	46.5	28.0	428	3	US-09-052-778-12
24	46.5	28.0	586	4	US-09-134-001C-4456
25	46	27.7	147	4	US-09-134-001C-3180
26	46	27.7	365	2	US-08-855-518-5
27	46	27.7	366	2	US-08-855-518-4

Sequence 2, Appli  
Sequence 14, Appl  
Sequence 16, Appl  
Sequence 18, Appl  
Sequence 20, Appl  
Sequence 22, Appl  
Sequence 24, Appl  
Sequence 26, Appl  
Sequence 28, Appl  
Sequence 30, Appl  
Sequence 32, Appl  
Sequence 34, Appl  
Sequence 36, Appl  
Sequence 38, Appl  
Sequence 40, Appl  
Sequence 42, Appl  
Sequence 44, Appl  
Sequence 46, Appl

28 46 27.7 553 2 US-08-943-087-2  
29 46 27.7 553 2 US-08-943-087-14  
30 46 27.7 553 2 US-08-943-087-16  
31 46 27.7 553 2 US-08-943-087-18  
32 46 27.7 553 2 US-08-943-087-20  
33 46 27.7 553 2 US-08-943-087-22  
34 46 27.7 553 2 US-08-943-087-24  
35 46 27.7 553 2 US-08-943-087-26  
36 46 27.7 553 2 US-08-943-087-28  
37 46 27.7 553 2 US-08-943-087-30  
38 46 27.7 553 2 US-08-943-087-32  
39 46 27.7 553 2 US-08-943-087-34  
40 46 27.7 553 2 US-08-943-087-36  
41 46 27.7 553 2 US-08-943-087-38  
42 46 27.7 553 2 US-08-943-087-40  
43 46 27.7 553 2 US-08-943-087-42  
44 46 27.7 553 2 US-08-943-087-44  
45 46 27.7 553 2 US-08-943-087-46

#### ALIGNMENTS

##### RESULT 1

US-09-052-778-14  
; Sequence 14, Application US/09052778A  
; Patent No. 6060590  
; GENERAL INFORMATION:  
; APPLICANT: Bryant, Peter J.  
; APPLICANT: Kawamura, Kazuo  
; TITLE OF INVENTION: CHITINASE RELATED PROTEINS AND METHODS  
; FILE REFERENCE: 07306/015001  
; CURRENT APPLICATION NUMBER: US/09/052,778A  
; CURRENT FILING DATE: 1998-03-31  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14  
; LENGTH: 441  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-052-778-14

Query Match 33.4%; Score 55.5; DB 3; Length 441;  
Best Local Similarity 48.3%; Pred. No. 8.3;  
Matches 14; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

Qy 2 LLLSLTSLVLLGSSWCGVPAITPALSY 30  
|||||: | : ||| | : | : | :  
Db 210 LLLSLTV-LPNVNSSWYDAPSIAPSLDF 237

##### RESULT 2

US-08-220-379B-7  
; Sequence 7, Application US/08220379B  
; Patent No. 5525708  
; GENERAL INFORMATION:  
; APPLICANT: No. 5525708ka, Karl  
; APPLICANT: Lobell, Robert B  
; TITLE OF INVENTION: STABILIZED DIMER OF KIT LIGAND  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Neave  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/220,379B  
; FILING DATE: 28-MAR-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haley Jr., James F.  
; REGISTRATION NUMBER: 27,794  
; REFERENCE/DOCKET NUMBER: CytoMed/2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-596-9000  
; TELEFAX: 212-596-9090  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 231 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; PUBLICATION INFORMATION:  
; AUTHORS: Lyman, et al.  
; JOURNAL: Cell  
; PAGES: 1157-1167  
; DATE: 1993  
; RELEVANT RESIDUES IN SEQ ID NO: 7: FROM 1 TO 231  
US-08-220-379B-7

Query Match 33.1%; Score 55; DB 1; Length 231;  
Best Local Similarity 61.1%; Pred. No. 4.7;  
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLGSSWG 18  
:|||||:|||||:  
Db 192 LLLLLLPLTLVLLAAWG 209

RESULT 3  
US-08-243-545-2  
; Sequence 2, Application US/08243545  
; Patent No. 5534512  
; GENERAL INFORMATION:  
; APPLICANT: Lyman, Stewart D.  
; APPLICANT: Beckmann, M. Patricia  
; TITLE OF INVENTION: Ligands for flt3/flk-2 Receptors  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Stephen L. Malaska, Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: Macintosh 7.0.1  
; SOFTWARE: Microsoft Word, Version #5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/243,545  
; FILING DATE: 11-MAY-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/162,407  
; FILING DATE: 03-DEC-1993  
; APPLICATION NUMBER: 08/111,758  
; FILING DATE: August 25, 1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/106,463  
; FILING DATE: August 12, 1993  
; CLASSIFICATION: 435

;; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/068,394  
; FILING DATE: May 24, 1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Malaska, Stephen L.  
; REGISTRATION NUMBER: 32,655  
; REFERENCE/DOCKET NUMBER: 2813-C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 587-0430  
; TELEFAX: (206) 233-0644  
; TELEX: 756822  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 231 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-243-545-2

Query Match 33.1%; Score 55; DB 1; Length 231;  
Best Local Similarity 61.1%; Pred. No. 4.7;  
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLGSSWG 18  
:|||||:|||||:  
Db 192 LLLLLLPLTLVLLAAWG 209

RESULT 4  
US-08-993-962-2  
; Sequence 2, Application US/08993962  
; Patent No. 5843423  
; GENERAL INFORMATION:  
; APPLICANT: Lyman, Stewart D.  
; APPLICANT: Beckmann, M. Patricia  
; TITLE OF INVENTION: Ligands for flt3/flk-2 Receptors  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Stephen L. Malaska, Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: Macintosh 7.0.1  
; SOFTWARE: Microsoft Word, Version #5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/993,962  
; FILING DATE: December 18, 1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/162,407  
; FILING DATE: December 3, 1993  
; APPLICATION NUMBER: 08/111,758  
; FILING DATE: August 25, 1993  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/106,463  
; FILING DATE: August 12, 1993  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/068,394  
; FILING DATE: May 24, 1993  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Malaska, Stephen L.  
; REGISTRATION NUMBER: 32,655  
; REFERENCE/DOCKET NUMBER: 2813-C  
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 587-0430  
TELEFAX: (206) 233-0644  
TELEX: 756822  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 231 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-993-962-2

Query Match 33.1%; Score 55; DB 2; Length 231;  
Best Local Similarity 61.1%; Pred. No. 4.7;  
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MLLSLTSLVLLGSSWG 18  
:|||||:|||||:  
Db 192 LLLLLPLTLVLLAAWG 209

## RESULT 5

US-09-160-841-2  
Sequence 2, Application US/09160841  
Patent No. 6190655  
GENERAL INFORMATION:  
APPLICANT: Lyman, Stewart D.  
APPLICANT: Beckmann, M. Patricia  
TITLE OF INVENTION: Ligands for flt3/flk-2 Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Stephen L. Malaska, Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: Washington  
COUNTRY: US  
ZIP: 98101

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh 7.0.1  
SOFTWARE: Microsoft Word, Version #5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/160,841  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/162,407  
FILING DATE: December 3, 1993  
APPLICATION NUMBER: 08/111,758  
FILING DATE: August 25, 1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/106,463  
FILING DATE: August 12, 1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/068,394  
FILING DATE: May 24, 1993

ATTORNEY/AGENT INFORMATION:  
NAME: Malaska, Stephen L.  
REGISTRATION NUMBER: 32,655  
REFERENCE/DOCKET NUMBER: 2813-C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 587-0430  
TELEFAX: (206) 233-0644  
TELEX: 756822  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 231 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-160-841-2

Query Match 33.1%; Score 55; DB 4; Length 231;

Best Local Similarity 61.1%; Pred. No. 4.7;  
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MLLSLTSLVLLGSSWG 18  
:|||||:|||||:  
Db 192 LLLLLPLTLVLLAAWG 209

## RESULT 6

PCT-US94-05365-2  
Sequence 2, Application PC/TUS9405365  
GENERAL INFORMATION:  
APPLICANT: Lyman, Stewart D.  
APPLICANT: Beckmann, M. Patricia  
TITLE OF INVENTION: Ligands for flt3/flk-2 Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Stephen L. Malaska, Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: Washington  
COUNTRY: US  
ZIP: 98101

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/05365  
FILING DATE: May 24, 1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: -to be assigned-  
FILING DATE: May 11, 1994

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/209,502  
FILING DATE: March 7, 1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/162,407  
FILING DATE: December 3, 1993  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/111,758  
FILING DATE: August 25, 1993

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/106,463  
FILING DATE: August 12, 1993  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/068,394  
FILING DATE: May 24, 1993  
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:  
NAME: Malaska, Stephen L.  
REGISTRATION NUMBER: 32,655  
REFERENCE/DOCKET NUMBER: 2813-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 587-0430  
TELEFAX: (206) 233-0644  
TELEX: 756822  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 231 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US94-05365-2

Query Match 33.1%; Score 55; DB 5; Length 231;

Best Local Similarity 61.1%; Pred. No. 4.7;  
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLGSSWG 18  
:|||||:|||||:|  
Db 192 LLLLLPLTLVLLAAWG 209

RESULT 7  
PCT-US95-03866-6  
; Sequence 6, Application PC/TUS9503866  
; GENERAL INFORMATION:  
; APPLICANT: Cytomed, Inc. (all states except US)  
; APPLICANT: Nocka, Karl (US only)  
; APPLICANT: Lobell, Robert B (US only)  
; TITLE OF INVENTION: STABILIZED DIMER OF KIT LIGAND AND  
; TITLE OF INVENTION: FLT-3/FLK-2 LIGAND  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Neave  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/03866  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/220,379  
; FILING DATE: 28-MAR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haley Jr, James F  
; REGISTRATION NUMBER: 27,794  
; REFERENCE/DOCKET NUMBER: Cytomed/2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-596-9000  
; TELEFAX: 212-596-9090  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 231 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; FEATURE:  
; NAME/KEY: Protein  
; LOCATION: 1..205  
PCT-US95-03866-6  
Query Match 33.1%; Score 55; DB 5; Length 231;  
Best Local Similarity 61.1%; Pred. No. 4.7;  
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MLLSLTSLVLGSSWG 18  
:|||||:|||||:|  
Db 192 LLLLLPLTLVLLAAWG 209  
RESULT 8  
US-08-270-584A-2  
; Sequence 2, Application US/08270584A  
; Patent No. 5710035  
; GENERAL INFORMATION:  
; APPLICANT: GREENE, ET AL.  
; TITLE OF INVENTION: Human Elastase IV

; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
; ADDRESSEE: CECCHI, STEWART & OLSTEIN  
; STREET: 6 BECKER FARM ROAD  
; CITY: ROSELAND  
; STATE: NEW JERSEY  
; COUNTRY: USA  
; ZIP: 07068  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 INCH DISKETTE  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WORD PERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/270,584A  
; FILING DATE: July 5, 1994  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: FERRARO, GREGORY D.  
; REGISTRATION NUMBER: 36,134  
; REFERENCE/DOCKET NUMBER: 325800-126  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-994-1700  
; TELEFAX: 201-994-1744  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 268 AMINO ACIDS  
; TYPE: AMINO ACID  
; STRANDEDNESS:  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
US-08-270-584A-2  
Query Match 33.1%; Score 55; DB 1; Length 268;  
Best Local Similarity 40.7%; Pred. No. 5.5;  
Matches 11; Conservative 5; Mismatches 11; Indels 0; Gaps 0;  
QY 3 LLSLTSLVLGSSWCGVPATPALS 29  
:|:|:|:|:|:|:|:|:|:|:|:|:|  
Db 1 MLGITVLAALLACASSCGVPFPNLS 27  
RESULT 9  
US-08-765-192-2  
; Sequence 2, Application US/08765192  
; Patent No. 5851814  
; GENERAL INFORMATION:  
; APPLICANT: Greene, John et al.  
; TITLE OF INVENTION: Human Elastase IV  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: MD  
; COUNTRY: US  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/765,192  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373

Db I MLGITVLAALLACASSCGVPSPFPNLS 27

APPLICANT: Stuart, Susan G.  
TITLE OF INVENTION: HUMAN EMRI-LIKE G PROTEIN COUPLED  
FILE REFERENCE: PF-0550 US  
CURRENT APPLICATION NUMBER: US/09/110,116  
CURRENT FILING DATE: 1998-07-02  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 4  
LENGTH: 344  
TYPE: PRT  
ORGANISM: HOMO SAPIENS  
FEATURE:  
OTHER INFORMATION: 2935597, GenBank  
US-09-110-116-4

Query Match 30.1%; Score 50; DB 3; Length 344;  
Best Local Similarity 37.0%; Pred. No. 35;  
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLGGSGGVPATP 27  
DB 250 MAFKATAQLFILGCTWGLGILQVGA 276

## RESULT 14

US-09-296-284-27  
Sequence 27, Application US/09296284A  
Patent No. 6204040

## GENERAL INFORMATION:

APPLICANT: Choi, Eui-Sung  
APPLICANT: Rhee, Sang-Ki  
APPLICANT: Lee, Eun-Hae  
TITLE OF INVENTION: Gluconobacter Suboxydans Sorbitol Dehydrogenase, Genes  
TITLE OF INVENTION: and Methods of Use Thereof

FILE REFERENCE: 1533.0870000  
CURRENT APPLICATION NUMBER: US/09/296,284A

CURRENT FILING DATE: 1999-04-22

NUMBER OF SEQ ID NOS: 87

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 27

LENGTH: 280

TYPE: PRT

ORGANISM: Gluconobacter suboxydans

US-09-296-284-27

Query Match 29.2%; Score 48.5; DB 4; Length 280;

Best Local Similarity 45.5%; Pred. No. 44;  
Matches 10; Conservative 3; Mismatches 8; Indels 1; Gaps 1;

QY 6 LTLSLVL-LGSSWGGVPATP 26  
DB 243 IRLALALARKWGAGVPKVLP 264

## RESULT 15

US-09-296-284-6  
Sequence 6, Application US/09296284A  
Patent No. 6204040

## GENERAL INFORMATION:

APPLICANT: Choi, Eui-Sung

APPLICANT: Rhee, Sang-Ki

APPLICANT: Lee, Eun-Hae

TITLE OF INVENTION: Gluconobacter Suboxydans Sorbitol Dehydrogenase, Genes  
TITLE OF INVENTION: and Methods of Use Thereof

FILE REFERENCE: 1533.0870000

CURRENT APPLICATION NUMBER: US/09/296,284A

CURRENT FILING DATE: 1999-04-22

NUMBER OF SEQ ID NOS: 87

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 6

LENGTH: 306

TYPE: PRT

ORGANISM: Gluconobacter suboxydans  
US-09-296-284-6

Query Match 29.2%; Score 48.5; DB 4; Length 306;

Best Local Similarity 45.5%; Pred. No. 48;  
Matches 10; Conservative 3; Mismatches 8; Indels 1; Gaps 1;

QY 6 LTLSLVL-LGSSWGGVPATP 26  
DB 269 IRLALALARKWGAGVPKVLP 290

Search completed: December 20, 2002, 15:20:38  
Job time : 2.86957 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 20, 2002, 15:12:13 ; Search time 5.08696 Seconds  
(without alignments)  
864.421 Million cell updates/sec

Title: US-09-856-319B-4\_COPY\_1\_33

Perfect score: 166

Sequence: 1 MLLSLTSLVLLGSGCGVPAITPALSINQR 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002.\*  
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	166	100.0	264	21	AA11711
2	150	90.4	264	21	AA11710
3	133	80.1	192	21	AA154077
4	68	41.0	44	21	AA164742
5	68	41.0	263	23	AA1982738
6	67	40.4	146	21	AA154191
7	59	35.5	195	20	AA111649
8	58	34.9	306	23	AA195588
9	58	34.9	312	22	AA171664
10	58	34.9	312	22	AA172192

11	58	34.9	312	22	AA172355	Human OR-like poly
12	55.5	33.4	441	21	AA107181	Drosophila melanog
13	55.5	33.4	441	22	AA10694	Drosophila melanog
14	55	33.1	123	20	AA11650	Human 5' EST seque
15	55	33.1	231	16	AA167540	Mouse flt-3 ligand
16	55	33.1	231	20	AA167768	Murine flt3-ligand
17	55	33.1	231	22	AA120186	Mouse flt-3 ligand
18	55	33.1	232	16	AA166177	Mouse Mtl10/Tl18
19	55	33.1	232	22	AA120189	Mouse flt-3 ligand
20	55	33.1	242	22	AA138137	Salmonella typhi c
21	55	33.1	268	17	AA190683	Human caldesmon c
22	55	33.1	268	17	AA188481	Human elastase IV
23	54	32.5	13	20	AA150212	Neutrophil-activat
24	54	32.5	84	22	AA131895	Novel human secret
25	52.5	31.6	760	22	AA166394	Human prostate ECG
26	52	31.3	55	22	AA167538	Amino acid sequenc
27	52	31.3	66	22	AA167539	Amino acid sequenc
28	52	31.3	222	20	AA137759	Amino acid sequenc
29	52	31.3	306	22	AA167544	Amino acid sequenc
30	52	31.3	319	21	AA136481	Fusion gene with h
31	52	31.3	319	22	AA167541	Amino acid sequenc
32	51	30.7	220	21	AA173464	Human secreted pro
33	51	30.7	306	20	AA141332	Human polypeptide
34	51	30.7	478	23	AA190142	Human protein sequ
35	51	30.7	484	22	AA193904	Herbicidally activ
36	51	30.7	826	23	AA191392	Human ORFX ORF1716
37	50.5	30.4	296	21	AA141952	Drosophila melanog
38	50	30.1	144	22	AA167458	Human prostate tum
39	50	30.1	162	20	AA173942	Rat camello 1 (Rcm
40	50	30.1	222	22	AA119992	Human cell surface
41	50	30.1	453	21	AA194340	Novel human diagno
42	50	30.1	517	22	AA14307	Human mature EGF-1
43	50	30.1	800	23	AA13609	Human EGF-like mol
44	50	30.1	823	23	AA13608	Nucleus specific
45	49.5	29.8	135	19	AA156733	

#### ALIGNMENTS

RESULT 1  
AA11711  
ID AA11711 standard; Protein: 264 AA.  
AC AA11711;  
XX 23-OCT-2000 (first entry)  
DT Mouse serine protease BSSP5 (mbSSP5) SEQ ID NO:4.  
XX BSSP5; serine protease; human; hbSSP5; mouse; mbSSP5; brain;  
XX diagnostic marker; antibody; transgenic animal; Alzheimer's disease;  
XX epilepsy; cancer; inflammation; infertility; pancreatitis;  
XX prostatic hypertrophy.  
OS Mus sp.  
PN WO200031243-A1.  
XX 02-JUN-2000.  
XX 19-NOV-1999; 99WO-JP06473.  
XX 20-NOV-1998; 98JP-0347806.  
XX (FUSO ) FUSO PHARM IND LTD.  
XX Uemura H, Okui A, Kominami K, Yamaguchi N, Mitsui S;  
XX WPI; 2000-400058/34.  
XX N-PSDB; AAA61734.  
XX Serine proteases BSSP5, useful in detecting homologs, mutants and

polymorphic variants as markers for diagnosis of e.g. Alzheimer's disease, epilepsy, cancer and inflammation, using blood, urine, pancreas or other tissues -

Claim 3: Page 55-56; 70pp; Japanese.

The invention relates to novel serine proteases designated BSSP5 (AAB11710-B11711), and to nucleic acids encoding them (AAA61733-A61734). The invention also relates to vectors and transformants comprising BSSP5 nucleic acids; transgenic animals in which the expression level of BSSP5 can be varied; and an mBSSP5 knockout mouse. The invention additionally encompasses anti-BSSP5 antibodies and methods of production of such antibodies, methods of BSSP5 detection using the antibodies, and the use of BSSP5 proteins or fragments as diagnostic markers for certain medical conditions, e.g., pancreatitis. A method for detecting pancreatitis comprising measuring BSSP5 concentration in the blood or urine, and a pancreatitis diagnostic agent containing an anti-BSSP5 antibody is also disclosed. Nucleotides encoding BSSP5 were initially isolated in a human brain cDNA library using degenerate PCR primers (AAA61744-A61745) based on conserved regions of serine proteases. The BSSP5 serine proteases and nucleotides encoding them are useful in detecting homologues, mutants and polymorphic variants in biological samples (e.g., blood, urine, brain, prostate gland, placenta, testis, pancreas and spleen) as diagnostic markers for conditions such as Alzheimer's disease, epilepsy, cancer, inflammation, infertility, pancreatitis and prostatic hypertrophy. Sequence AAB11710 represents human BSSP5 (hBSSP5), and sequence AAB11711 represents murine BSSP5 (mBSSP5).

Sequence 264 AA;

Query Match 100.0%; Score 166; DB 21; Length 264;  
Best Local Similarity 100.0%; Pred. No. 1e-13; Mismatches 0; Indels 0; Gaps 0;  
Matches 33; Conservative 0;

QY 1 MLLSLTSLVLGGSGGCGVPAITPALSYNQR 33  
|||||  
DB 1 MLLSLTSLVLGGSGGCGVPAITPALSYNQR 33

RESULT 2

AAB11710  
ID AAB11710 standard; Protein; 264 AA.

AC AAB11710;

DT 23-OCT-2000 (first entry)

DE Human serine protease BSSP5 (hBSSP5) SEQ ID NO:2.

DE BSSP5; serine protease; human; hBSSP5; mouse; mBSSP5; brain; diagnostic marker; antibody; transgenic animal; Alzheimer's disease; epilepsy; cancer; inflammation; infertility; pancreatitis; prostatic hypertrophy.

OS Homo sapiens.

PN WO200031243-A1.

PD 02-JUN-2000.

PF 19-NOV-1999; 99WO-JP06473.

PR 20-NOV-1998; 98JP-0347806.

PA (FUSO ) FUSO PHARM IND LTD.

PI Uemura H, Okui A, Kominami K, Yamaguchi N, Mitsui S;

DR WPI: 2000-400058/34.

DR N-PSDB; AAA61733.

XX Serine proteases BSSP5, useful in detecting homologs, mutants and

polymorphic variants as markers for diagnosis of e.g. Alzheimer's disease, epilepsy, cancer and inflammation, using blood, urine, pancreas or other tissues -

Claim 1: Page 51-52; 70pp; Japanese.

The invention relates to novel serine proteases designated BSSP5 (AAB11710-B11711), and to nucleic acids encoding them (AAA61733-A61734). The invention also relates to vectors and transformants comprising BSSP5 nucleic acids; transgenic animals in which the expression level of BSSP5 can be varied; and an mBSSP5 knockout mouse. The invention additionally encompasses anti-BSSP5 antibodies and methods of production of such antibodies, methods of BSSP5 detection using the antibodies, and the use of BSSP5 proteins or fragments as diagnostic markers for certain medical conditions, e.g., pancreatitis. A method for detecting pancreatitis comprising measuring BSSP5 concentration in the blood or urine, and a pancreatitis diagnostic agent containing an anti-BSSP5 antibody is also disclosed. Nucleotides encoding BSSP5 were initially isolated in a human brain cDNA library using degenerate PCR primers (AAA61744-A61745) based on conserved regions of serine proteases. The BSSP5 serine proteases and nucleotides encoding them are useful in detecting homologues, mutants and polymorphic variants in biological samples (e.g., blood, urine, brain, prostate gland, placenta, testis, pancreas and spleen) as diagnostic markers for conditions such as Alzheimer's disease, epilepsy, cancer, inflammation, infertility, pancreatitis and prostatic hypertrophy. Sequence AAB11710 represents human BSSP5 (hBSSP5), and sequence AAB11711 represents murine BSSP5 (mBSSP5).

Sequence 264 AA;

Query Match 90.4%; Score 150; DB 21; Length 264;  
Best Local Similarity 87.9%; Pred. No. 1.3e-11;  
Matches 29; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLGGSGGCGVPAITPALSYNQR 33  
|||||  
DB 1 MLLSLTSLVLGGSGGCGVPAITPALSYNQR 33

RESULT 3

AAB54077  
ID AAB54077 standard; Protein; 192 AA.

AC AAB54077;

DT 09-MAR-2001 (first entry)

DE Human pancreatic cancer antigen protein sequence SEQ ID NO:529.

DE Human; pancreas; pancreatic cancer; pancreatic cancer antigen; detection; diagnosis; identification; cytostatic; neuroprotective; nontropic; immunomodulatory; relaxant; contraceptive; gynaecological; antinflammatory; cardiac; gene therapy; chromosome mapping; linkage analysis; tissue identification; tissue typing; forensic; neural; immune system; muscular; reproductive; gastrointestinal; pulmonary; cardiovascular; renal; proliferative.

OS Homo sapiens.

PN WO200055320-A1.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000WO-US05989.

PR 12-MAR-1999; 99US-0124270.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM;

XX WPI: 2000-579444/54.



XX Nucleic acids encoding novel human proteases, useful for  
 PT treating diseases and disorders such as cancers, immune-related  
 PT diseases and disorders, cardiovascular disease (e.g. restenosis) and  
 PT inflammatory disorders -  
 XX  
 XX Claim 6; Fig 2M; 313pp; English.  
 XX  
 XX The present invention relates to the isolation of novel human  
 CC proteases, and the nucleic acids encoding them. The sequences of  
 CC the invention are useful for treating diseases and disorders such as  
 CC cancers (e.g. breast, colon, lung), immune-related diseases and disorders  
 CC (e.g. inflammatory diseases and asthma), cardiovascular diseases  
 CC (e.g. restenosis and coronary thrombosis), brain or neuronal-associated  
 CC diseases, metabolic disorders (e.g. diabetes, obesity), inflammatory  
 CC disorders (e.g. rheumatoid arthritis and psoriasis), central or  
 CC peripheral nervous system diseases, migraines, pain, sexual dysfunction,  
 CC mood disorders, attention disorders, cognition disorders, hypertension,  
 CC hypertension, psychotic disorders, neurological disorders  
 CC (e.g. Alzheimer's disease, Parkinson's disease) and dyskinesias.  
 CC The nucleic acids and polypeptides are also useful for treating viral  
 CC infections caused by human immunodeficiency virus (HIV), and non-viral  
 CC infections such as ocular disease (e.g. glaucoma) and macular  
 CC degeneration. AAU82702-AAU82760 represent the novel human proteases of  
 CC the invention.  
 XX  
 XX Sequence 263 AA;

Query Match 41.0%; Score 68; DB 23; Length 263;  
 Best Local Similarity 51.7%; Pred. No. 0.63;  
 Matches 15; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLLGSSWCGVPAITPALS 29  
 Db 1 MASLWLLCSFLVGAFCGVPPIHPVLS 29

RESULT 6  
 AAB54191  
 ID AAB54191 standard; Protein; 146 AA.  
 XX  
 AC AAB54191;  
 XX  
 XX 09-MAR-2001 (first entry)  
 XX  
 XX Human pancreatic cancer antigen protein sequence SEQ ID NO:643.  
 XX  
 XX Human; pancreas; pancreatic cancer; pancreatic cancer antigen;  
 KW detection; diagnosis; identification; cytostatic; neuroprotective;  
 KW nontropic; immunomodulatory; relaxant; contraceptive; gynaecological;  
 KW antiinflammatory; cardiant; gene therapy; chromosome mapping;  
 KW linkage analysis; tissue identification; tissue typing; forensic;  
 KW neural; immune system; muscular; reproductive; gastrointestinal;  
 KW pulmonary; cardiovascular; renal; proliferative.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO2000055320-A1.  
 XX  
 XX 21-SEP-2000.  
 XX  
 XX 08-MAR-2000; 2000WO-US05989.  
 XX  
 XX 12-MAR-1999; 99US-0124270.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Rosen CA, Ruben SM;  
 XX  
 XX WPI; 2000-579444/54.  
 XX  
 XX N-PSDB; AAC98956.  
 XX  
 XX New nucleic acid that is a pancreatic cancer antigen for preventing,

PT treating, or ameliorating a medical condition, particular pancreatic  
 PT cancer, or for use in assays for diagnosing a pathological condition -  
 XX  
 XX Claim 11; Page 1081; 1379pp; English.  
 XX  
 XX AAC98773 to AAC99231 encode the human pancreatic cancer associated  
 CC proteins, called pancreatic cancer antigens, given in AAB54008 to  
 CC AAB54466. The human pancreatic cancer antigens have cytostatic,  
 CC neuroprotective, nontropic, immunomodulatory, relaxant, contraceptive,  
 CC gynaecological, cardiant and antiinflammatory activities, and can be used  
 CC in gene therapy. The polynucleotide and proteins can be used for  
 CC preventing, treating, or ameliorating a medical condition or in assays  
 CC for diagnosing a pathological condition or a susceptibility to one in a  
 CC subject. Binding partners to the proteins and the activity of the  
 CC proteins can be identified. The pancreatic cancer antigens can be used to  
 CC detect, treat or prevent pancreatic disorders, especially cancer.  
 CC Agonists and antagonists to the antigens can be screened for. The  
 CC pancreatic cancer antigen polynucleotides can be used to design nucleic  
 CC acid hybridisation probes that can be used in chromosome mapping, linkage  
 CC analysis, tissue identification and/or typing and a variety of forensic  
 CC and diagnostic methods. The proteins can be used to generate antibodies  
 CC which are used to purify, detect and target the polypeptides, including  
 CC both in vivo and in vitro diagnostic and therapeutic methods. The  
 CC proteins can be used to treat or prevent neural, immune system, muscular,  
 CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or  
 CC proliferative disorders: AAC99232 to AAC99240 and AAB54467 represent  
 CC sequences used in the exemplification of the present invention.  
 XX

SQ Sequence 146 AA;

Query Match 40.4%; Score 67; DB 21; Length 146;  
 Best Local Similarity 66.7%; Pred. No. 0.45;  
 Matches 12; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 12 LLGSSWCGVPAITPALS 29  
 Db 16 LVGAFCGVPPIHPVLS 33

RESULT 7  
 AAY11649  
 ID AAY11649 standard; Protein; 195 AA.  
 XX  
 AC AAY11649;  
 XX  
 XX 16-JUN-1999 (first entry)  
 XX  
 XX Human 5' EST secreted protein SEQ ID NO:301.  
 DE  
 XX Human; secreted protein; EST; expressed sequence tag; diagnosis;  
 KW forensic; gene therapy; chromosome mapping; signal peptide;  
 KW upstream regulatory sequence; cytokine activity; cell proliferation;  
 KW differentiation; haematopoiesis regulation; tissue growth regulation;  
 KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;  
 KW thrombolytic; anti-inflammatory; tumour inhibition.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO9906439-A2.  
 XX  
 XX 11-FEB-1999.  
 XX  
 XX 31-JUL-1998; 98WO-IB01233.  
 XX  
 XX 01-AUG-1997; 97US-0904468.  
 XX  
 XX (GEST ) GENSET.  
 XX  
 XX Duclert A, Dumas Milne Edwards J, Lacroix B;  
 XX  
 XX WPI; 1999-153700/13.  
 XX  
 XX N-PSDB; AAX40367.  
 XX

PT New nucleic acids encoding human secreted proteins - obtained from  
PT cDNA libraries derived from liver, lung, large intestine, colon,  
PT thyroid and pancreas tissue  
XX  
XX Claim 27; Page 378-379; 398pp; English.  
XX  
CC AAX40251 to AAX40397 represent 5' expressed sequence tags (ESTs) for  
CC human secreted proteins, and encode the proteins given in AAY11533 to  
CC AAY11679, respectively. The proteins given represent the signal peptide  
CC and an N-terminal fragment of a secreted protein. The nucleic acid  
CC sequences can be used for producing secreted human gene products. They  
CC can also be used to develop products for diagnosis and therapy. The  
CC proteins obtained may have cytokine activity, cell  
CC proliferation/differentiation activity, haematopoiesis regulating  
CC activity, tissue growth regulating activity, reproductive hormone  
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and  
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory  
CC activity, tumour inhibition activity or other activities. The products  
CC can be used in forensic, gene therapy and chromosome mapping procedures.  
CC The sequences can also be used for obtaining corresponding promoter  
CC sequences. The nucleic acids encoding the signal peptide can be used for  
CC directing extracellular secretion of a polypeptide or the insertion of a  
CC polypeptide into a membrane, or importing a polypeptide into a cell.  
XX  
XX Sequence 195 AA;  
SQ  
Query Match 35.5%; Score 59; DB 20; Length 195;  
Best Local Similarity 44.4%; Pred. No. 6.8;  
Matches 12; Conservative 5; Mismatches 10; Indels 0; Gaps 0;  
Oy 3 LLSLTSLVLGGSGGCGVPAITPALS 29  
Db 1 MLSITVLAALLACASSCGVSPFPNLS 27  
  
RESULT 8  
AAU95588  
ID AAU95588 standard; Protein: 306 AA.  
XX  
XX AAU95588;  
AC  
DT 02-JUL-2002 (first entry)  
XX  
XX Human olfactory and pheromone G protein-coupled receptor #75.  
XX  
XX Human; olfactory and pheromone G protein coupled; receptor;  
KW GPCR; tranquilizer; antidepressant; neuroleptic; endocrine; anabolic;  
KW anorectic; taste; fragrance; food additive; cosmetic; cell migration;  
KW sterility; psychotic disorder; neurological disorder; anxiety;  
KW schizophrenia; manic depression; depression; axonal growth;  
KW menstrual cycle; appetite sexual motivation; sexual attraction;  
KW aggression.  
XX  
XX Homo sapiens.  
OS  
XX W0200224726-A2.  
PN  
XX 28-MAR-2002.  
PD  
XX  
XX 21-SEP-2001; 2001WO-BE00162.  
PF  
XX 22-SEP-2000; 2000EP-0870211.  
PR  
XX (CHEM-) CHEMCOM SA.  
PA  
XX Veithen A;  
PI  
XX WPI; 2002-330013/36.  
DR  
XX N-PSDB; ABK68475.  
DR  
XX Novel pheromone G-protein coupled receptor and receptor-derived  
PT agonists, antagonists or inhibitors useful in food or cosmetic products  
PT or in the treatment or prevention of neurological disorders such as

PT anxiety and schizophrenia -  
XX  
XX Disclosure; Page 238-239; 833pp; English.  
XX  
CC The invention relates to olfactory and pheromone G-protein coupled  
CC receptor (GPCR) or a protein 95% identical to the GPCR, a specific active  
CC portion and its encoding polynucleotide. Also included are an agonist,  
CC antagonist or inhibitor of the GPCR or the polynucleotide, a vector  
CC comprising the polynucleotide, a cell transformed by the vector, a non-  
CC human mammal comprising a partial or total deletion of the polynucleotide  
CC encoding the receptor and screening (detection and possibly, recovering)  
CC of compounds which are known or not known to be agonist, antagonists or  
CC inhibitors of natural compounds to the GPCR. The receptor-derived  
CC agonists, antagonists, inhibitors or compounds are used as an  
CC improvement, elimination or substitution of an existing taste and/or a  
CC fragrance of (or in) the food and/or cosmetic products. They can also be  
CC used in the preparation of medicament in the treatment and/or prevention  
CC of a mammalian disorder, such as cell migration, sterility, psychotic and  
CC neurological disorders, including anxiety, schizophrenia, manic  
CC depression, depression, for promoting axonal growth, nerve cell  
CC connection and nerve regeneration for modulating male and female  
CC endocrine functions, hormone production and the menstrual cycle, for the  
CC prevention or the treatment by stimulation of several mammalian  
CC behaviours, such as stimulation or suppression of appetite, sexual  
CC motivation, sexual attraction, aggression and for promoting or  
CC suppressing chemical communication between organisms. The present  
CC sequence is a human olfactory and pheromone GPCR protein sequence.  
XX  
XX Sequence 306 AA;  
SQ  
Query Match 34.9%; Score 58; DB 23; Length 306;  
Best Local Similarity 60.0%; Pred. No. 15;  
Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
Oy 2 LLSLTSLVLGGSGGCGV 21  
Db 127 LLSHRVCLAGSAWACGV 146  
  
RESULT 9  
AAG71664  
ID AAG71664 standard; Protein: 312 AA.  
XX  
XX AAG71664;  
AC  
DT 30-JUL-2001 (first entry)  
XX  
XX Human olfactory receptor polypeptide, SEQ ID NO: 1345.  
DE Human; olfactory receptor; OR; primary scent determination;  
XX secondary scent determination; polypeptide library; odour receptor;  
KW scent profile; scent fingerprint; scent representation.  
KW  
XX Homo sapiens.  
OS  
XX W0200127158-A2.  
PN  
XX 19-APR-2001.  
PD  
XX 06-OCT-2000; 2000WO-US27582.  
PF  
XX 08-OCT-1999; 99US-0158615.  
PR  
XX 24-FEB-2000; 2000US-0184809.  
PR  
XX (DIGI-) DIGISCENTS.  
PA --(YEDA) YEDA RES & DEV CO LTD.  
XX  
XX Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;  
PI  
XX WPI; 2001-290713/30.  
DR  
XX New polynucleotides which encode polypeptides involved in olfactory  
PT sensation for identifying olfactory agonists and antagonists -

xx Claim 11; Page 830-831; 1857pp; English.

xx The present sequence is an olfactory receptor which is encoded by

cc one of a number of novel polynucleotides. The polynucleotides can be

cc used in screening for olfactory agonists and antagonists. The methods

cc allow for the determination of primary scents and the identification

cc of the odour receptors used to detect these primary scents. The methods

cc also enable determination of secondary scents and the identification of

cc combinations of odour receptors that are involved in detecting such

cc secondary scents. This enables the construction of a scent representation

cc (also called a scent fingerprint or scent profile), which may be used to

cc re-create and edit scents. Libraries of olfactory receptors are useful

cc for determining the interaction pattern of a composition with the

cc receptors, and can be used for determining differences in the olfactory

cc faculties of different individuals.

xx Sequence 312 AA;

QY 2 LLLSLTSLVLGSSWGCGV 21  
|||| : |||: |||

Db 133 LLLSHRVCLQAGSAWACGV 152

RESULT 11

AAG72192

ID AAG72192 standard; Protein; 312 AA.

AC AAG72192;

XX 30-JUL-2001 (first entry)

DT Human olfactory receptor polypeptide, SEQ ID NO: 1873.

DE Human; olfactory receptor; OR; primary scent determination;

XX secondary scent determination; polypeptide library; odour receptor;

KW scent profile; scent fingerprint; scent representation.

XX Homo sapiens.

OS WO200127158-A2.

PN 19-APR-2001.

PD 06-OCT-2000; 2000WO-US27582.

PF 08-OCT-1999; 99US-0158615.

XX 24-FEB-2000; 2000US-0184809.

XX (DIGI-) DIGISCENTS.

PA (YEDA ) YEDA RES & DEV CO LTD.

XX Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;

PI WPI; 2001-290713/30.

XX New polynucleotides which encode polypeptides involved in olfactory

PT sensation for identifying olfactory agonists and antagonists -

XX Claim 11; Page 1240; 1857pp; English.

xx The present sequence is an olfactory receptor which is encoded by

cc one of a number of novel polynucleotides. The polynucleotides can be

cc used in screening for olfactory agonists and antagonists. The methods

cc allow for the determination of primary scents and the identification

cc of the odour receptors used to detect these primary scents. The methods

cc also enable determination of secondary scents and the identification of

cc combinations of odour receptors that are involved in detecting such

cc secondary scents. This enables the construction of a scent representation

cc

QY 2 LLLSLTSLVLGSSWGCGV 21  
|||| : |||: |||

Db 133 LLLSHRVCLQAGSAWACGV 152

RESULT 11

AAG72355

ID AAG72355 standard; Protein; 312 AA.

AC AAG72355;

XX 30-JUL-2001 (first entry)

DT Human OR-like polypeptide query sequence, SEQ ID NO: 2036.

DE Human; olfactory receptor; OR; primary scent determination;

XX secondary scent determination; polypeptide library; odour receptor;

KW scent profile; scent fingerprint; scent representation.

XX Homo sapiens.

OS WO200127158-A2.

PN 19-APR-2001.

PD 06-OCT-2000; 2000WO-US27582.

PF 08-OCT-1999; 99US-0158615.

XX 24-FEB-2000; 2000US-0184809.

XX (DIGI-) DIGISCENTS.

PA (YEDA ) YEDA RES & DEV CO LTD.

XX Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;

PI WPI; 2001-290713/30.

XX New polynucleotides which encode polypeptides involved in olfactory

PT sensation for identifying olfactory agonists and antagonists -

XX Example 6; Page 1364; 1857pp; English.

xx The present sequence is a polypeptide encoded by one of 344 newly mined

cc human genes. It was used as a query sequence in a database search of

cc olfactory receptor (OR)-like sequences. The invention relates to isolated

cc polynucleotides encoding polypeptides involved in olfactory sensation.

cc The polynucleotides can be used in screening for olfactory agonists and

cc antagonists. The methods allow for the determination of primary scents

cc and the identification of the odour receptors used to detect these

cc primary scents. The methods also enable determination of secondary scents

cc and the identification of combinations of odour receptors that are

cc involved in detecting such secondary scents. This enables the

cc construction of a scent representation (also called a scent fingerprint

cc or scent profile), which may be used to re-create and edit scents.

cc Libraries of olfactory receptors are useful for determining the

cc interaction pattern of a composition with the receptors, and can be

cc used for determining differences in the olfactory faculties of different

cc individuals.

xx Sequence 312 AA;

Query Match 34.9%; Score 58; DB 22; Length 312;  
Best Local Similarity 60.0%; Pred. No. 15;  
Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
QY 2 LLLSLTSLVLLGSSWGCV 21  
||||| : | | | | |  
Db 133 LLLSHRVCLQAGSAMACGV 152

RESULT 12  
AAB07181  
ID AAB07181 standard; Protein; 441 AA.  
XX  
AC AAB07181;  
DT 22-NOV-2000 (first entry)  
DE Drosophila melanogaster imaginal disc growth factor 3.  
XX  
KW Imaginal disc growth factor 3; IDGF3; chitinase related protein;  
KW CHRP; tissue growth; wound healing; bone repair; cartilage repair;  
KW angiogenesis; meat production; milk production; cancer; gene therapy.  
XX  
OS Drosophila melanogaster.  
XX  
US06060590-A.  
PN  
XX  
PD 09-MAY-2000.  
XX  
PF 31-MAR-1998; 98US-0052778.  
XX  
PR 31-MAR-1998; 98US-0052778.  
XX  
PA (REGC ) UNIV CALIFORNIA.  
XX  
PI Bryant PJ, Kawamura K;  
XX  
WPI; 2000-349702/30.  
XX  
PT New imaginal disc growth factor polypeptide, useful for healing wounds,  
PT promoting tissue growth, modulating female reproductive tract functions  
PT and treating cell proliferative disorders such as cancer and metastasis  
PT .  
XX  
PS Disclosure; Fig 4; 40pp; English.

XX  
XX The present sequence is the protein sequence for the Drosophila  
CC melanogaster imaginal disc growth factor 3 (IDGF3). The sequence encoding  
CC this protein was used to isolate the IDGF4 coding sequence. IDGF4 is a  
CC member of the chitinase related proteins (CHRP) and is involved in the  
CC promotion of cell growth, motility and morphogenesis. The gene and  
CC protein are expected to have mammalian homologues. They can be used in  
CC the treatment of cancer, wound healing, tissue regeneration following  
CC arthritis, osteoporosis, other skeletal disorders and burns, for  
CC revitalising scar tissue resulting from surgical procedures, irradiation,  
CC laceration, toxic chemicals, viral or bacterial infection or burns, to  
CC promote tissue growth during tissue engineering, for example tissues for  
CC skin graft replacements and bone regrowth, and to modulate the function  
CC of the female reproductive tract. In addition, they can also be used to  
CC increase meat, egg, sperm and milk production in animals. One possible  
CC method of use is by gene therapy.  
XX  
SQ Sequence 441 AA;

Query Match 33.4%; Score 55.5; DB 21; Length 441;  
Best Local Similarity 48.3%; Pred. No. 47;  
Matches 14; Conservative 5; Mismatches 9; Indels 1; Gaps 1;  
QY 2 LLLSLTSLVLLGSSWGCVPAITPALS 30  
||||| : | | | | |  
Db 210 LLLSLTV-LPNVNSSWYDAPSIAPSLDF 237

RESULT 13  
ABB60694  
ID ABB60694 standard; Protein; 441 AA.  
XX  
AC ABB60694;  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 8874.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
WO200171042-A2.  
PN  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US09231.  
XX  
PR 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
DR N-PSDB; ABL04797.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
PS Disclosure; SEQ ID NO 8874; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (ABB57737-ABB72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 441 AA;

Query Match 33.4%; Score 55.5; DB 22; Length 441;  
Best Local Similarity 48.3%; Pred. No. 47;  
Matches 14; Conservative 5; Mismatches 9; Indels 1; Gaps 1;  
QY 2 LLLSLTSLVLLGSSWGCVPAITPALS 30  
||||| : | | | | |  
Db 210 LLLSLTV-LPNVNSSWYDAPSIAPSLDF 237

RESULT 14  
AAV11650  
ID AAV11650 standard; Protein; 123 AA.  
XX  
AC AAV11650;  
DT 16-JUN-1999 (first entry)  
XX  
DE Human 5' EST secreted protein SEQ ID NO:302.  
XX  
KW Human; secreted protein; EST; expressed sequence tag; diagnosis;  
KW forensic; gene therapy; chromosome mapping; signal peptide;  
KW upstream regulatory sequence; cytokine activity; cell proliferation;

KW differentiation; haematopoiesis regulation; tissue growth regulation;  
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;  
KW thrombolytic; anti-inflammatory; tumour inhibition.

OS Homo sapiens.

XX WO9906439-A2.

XX 11-FEB-1999.

XX 31-JUL-1998; 98WO-IB01233.

XX 01-AUG-1997; 97US-0904468.

XX (GEST ) GENSET.

XX Duclert A, Dumas Milne Edwards J, Lacroix B;

XX WPI; 1999-153700/13.

DR N-PSDB; AAX40368.

XX New nucleic acids encoding human secreted proteins - obtained from  
PT cDNA libraries derived from liver, lung, large intestine, colon,  
PT thyroid and pancreas tissue

PS Claim 27; Page 379-380; 398pp; English.

CC AAX40251 to AAX40397 represent 5' expressed sequence tags (ESTs) for  
CC human secreted proteins, and encode the proteins given in AAY11533 to  
CC AAY11679, respectively. The proteins given represent the signal peptide  
CC and an N-terminal fragment of a secreted protein. The nucleic acid  
CC sequences can be used for producing secreted human gene products. They  
CC can also be used to develop products for diagnosis and therapy. The  
CC proteins obtained may have cytokine activity, cell  
CC proliferation/differentiation activity, haematopoiesis regulating  
CC activity, tissue growth regulating activity, reproductive hormone  
CC regulating activity, chemotactic/ chemokinetic activity, haemostatic and  
CC thrombolytic activity, receptor/ ligand activity, anti-inflammatory  
CC activity, tumour inhibition activity or other activities. The products  
CC can be used in forensic, gene therapy and chromosome mapping procedures.  
CC The sequences can also be used for obtaining corresponding promoter  
CC sequences. The nucleic acids encoding the signal peptide can be used for  
CC directing extracellular secretion of a polypeptide or the insertion of a  
CC polypeptide into a membrane, or importing a polypeptide into a cell.

XX Sequence 123 AA;

Query Match 33.1%; Score 55; DB 20; Length 123;

Best Local Similarity 40.7%; Pred. No. 14;

Matches 11; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 3 LLSLTSLVLLGSGGCVPAITPALS 29

DB 1 MLGITVLAALLACASSCGVFPFNLS 27

RESULT 15

AAR67540

ID AAR67540 standard; Protein; 231 AA.

XX AAR67540;

XX 05-AUG-1995 (first entry)

XX Mouse flt-3 ligand.

XX Flt-3 ligand; flt3-L; anemia; cancer; AIDS; gene therapy.

XX Mus sp.

XX Key

XX Location/Qualifiers

XX 1..27

XX /label= Sig\_peptide

FT

FT Domain

28..188

/label= Extracellular\_domain

FT Domain

189..211

/label= Transmembrane\_domain

FT Domain

212..231

/label= Cytoplasmic\_domain

XX EP627487-A.

PN 07-DEC-1994.

XX 19-MAY-1994; 94EP-0303575.

XX 24-MAY-1993; 93US-0068394.

PR 12-AUG-1993; 93US-0106463.

PR 25-AUG-1993; 93US-0111758.

PR 03-DEC-1993; 93US-0162407.

PR 07-MAR-1994; 94US-0209502.

PR 11-MAY-1994; 94US-0243545.

XX (IMMV ) IMMUNEX CORP.

XX Beckmann MP, Lyman SD;

PI WPI; 1995-008071/02.

XX N-PSDB; AAQ79076.

DR Isolated ligands for flt 3 receptors - useful for treating

XX anaemia, AIDS and various cancers

XX Disclosure; Page 25-27; 33pp; English.

CC cDNA encoding mouse flt3-ligand (flt3-L) was isolated from a

CC cDNA library of T-cell line P7B-0.3A4 in CV-1/EBNA-1 cells

CC using a slide autoradiography method. Flt3-L stimulates

CC production of progenitor and stem cells, and can be used e.g.

CC in gene therapy protocols.

XX Sequence 231 AA;

Query Match 33.1%; Score 55; DB 16; Length 231;

Best Local Similarity 61.1%; Pred. No. 27;

Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLLSLTSLVLLGSSWG 18

DB 192 LLLLLLPTLVLLAAWG 209

Search completed: December 20, 2002, 15:16:42

Job time : 6.08696 secs